MASS SPECTROMETRY-BASED CHARACTERIZATION, QUANTITATION, AND REPAIR INVESTIGATIONS OF COMPLEX DNA LESIONS

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF THE
UNIVERSITY OF MINNESOTA
BY

Arnold Groehler IV

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Dr. Natalia Y. Tretyakova, Advisor

March 2018

Acknowledgements

As my time as a graduate student comes to an end, I must express my sincere thanks and gratitude to all the people who helped me during the process. First and foremost, I would like to thank my advisor, Dr. Natalia Tretyakova, for her constant support and guidance over the last six years. As an aspiring professor, Dr. Tretyakova's dedication to helping her students succeed and produce quality science was inspiring. I would also like to thank Dr. Colin Campbell for his mentorship in biology and constant support during my graduate career. I am also grateful to Dr. Daniel A. Harki and Dr. Rodney L. Johnson for agreeing to serve as my thesis committee members and for their valuable feedback and suggestions.

I would also like to express my thanks to all the Tretyakova and Campbell group members for their mentorship, contributions, and discussions. Dr. Susith Wickramaratne, Dr. Dewakar Sangaraju, Dr. Srikanth Kotapati, Dr. Teshome Gherezghiher, and Dr. Delshanee Kotandeniya for their mentorship during the first years of my graduate work. Abby Hoffman, Joshua Schmidt, Dominic Najjar, Lisa Chesner, Dr. Suresh Pujari, and Daeyoon Park, whose contributions and support were crucial to completing my projects. And all current and past members of the Tretyakova laboratory for their helpful suggestions and contributions throughout the past six years.

Furthermore, I would like to thank all the collaborators who contributed to my projects. My sincere thanks to Dr. Yuichi Machida, Dr. Reeja Maskey, Dr. Mary Garry, Qinglu Li, Stefan Kren, Maggie Robledo-Villafane, Dr. Deborah Ferrington, and Dr. Ashis Basu for performing experiments and/or generous donation of tissues for my thesis projects. I am also very thankful to Brock Matter, Dr. Peter Villalta, Xun Ming, and

Yingchun Zhao for their help and expertise in mass spectrometry method development and troubleshooting. Furthermore, I would like to thank Robert Carlson for his help with preparing figures for my publications, presentations, and thesis.

Last but not least, I would like to thank my family and friends, especially my incoming classmates for their unconditional support and encouragement. I could not have achieved such a great accomplishment without all of you in my life.

Abstract

DNA is constantly under the threat of damage by various endogenous and exogenous agents, leading to the structural modification of nucleobases (DNA adducts). These DNA adducts can range from smaller nucleoside monoadducts and exocyclic adducts, to the helix distorting and super-bulky DNA-DNA cross-links and DNA-protein cross-links. If not repaired, DNA adducts can inhibit crucial biological processes such as DNA replication, leading to adverse consequences such as mutagenesis and carcinogenesis. Therefore, understanding the atomic connectivity, extent of formation, and repair of DNA adducts is crucial to fully elucidating the biological consequences of the adduct.

DNA-protein cross-links (DPCs) are ubiquitous, super-bulky DNA lesions that form when proteins become irreversibly trapped on chromosomal DNA. The structural complexity of cross-linking and the diversity of proteins susceptible to DPC formation represents significant challenges to studying the biological consequence of these adducts. In the first part of the thesis, we identified the protein constituents, structural characterized and quantified, and investigated the repair mechanism of *bis*-electrophile (**Chapter 2**) and reactive oxygen species (ROS, **Chapters 3** and **4**)-induced DPCs.

In **Chapter 2**, we investigated DPC formation after exposure to N,N-*bis*-(2-chloroethyl)-phosphorodiamidic acid (phosphoramide mustard, PM) and N,N-*bis*-(2-chloroethyl)-ethylamine (nornitrogen mustard, NOR), the two biologically active metabolites of the antitumor agent cyclophosphamide. A mass spectrometry-based proteomics approach was employed to characterize the protein constituents of PM- and NOR-mediated DNA-protein cross-linking in human fibrosarcoma (HT1080) cells. HPLC-

ESI^+-MS/MS analysis of proteolytic digests of DPC-containing DNA from NOR-treated cells revealed a concentration-dependent formation of *N*-[2-[cysteinyl]ethyl]-N-[2-(guan-7-yl)ethyl]amine (Cys-NOR-N7G) conjugates, confirming that it cross-links cysteine thiols of proteins to the N-7 position of guanines in DNA. A sensitive and accurate Cys-NOR-N7G isotope dilution tandem mass spectrometry assay was developed to quantify PM-induced DPC formation and repair in mammalian cells proficient or deficient in a DNA repair pathway.

In **Chapters 3**, we employed the model of left anterior descending artery ligation/reperfusion surgery in rat to show that ischemia/reperfusion injury is associated with the formation of hydroxyl radical-induced DNA-protein cross-links (DPCs) in cardiomyocytes. Mass spectrometry based experiments revealed that these conjugates were formed by a free radical mechanism and involved thymidine residues of DNA and tyrosine side chains of proteins (dT-Tyr). Quantitative proteomics experiments utilizing Tandem mass tags (TMT) revealed that radical-induced DPC formation increase after LAD-ligation/reperfusion compared to the control sham surgery. Using the developed dT-Tyr nanoLC-ESI^+-MS/MS assay, we investigated the role of the metalloprotease Spartan (*SPRTN*) in the repair of radical-induced DPCs (**Chapter 4**). Analysis of the brain, liver, heart, and kidneys of wild type ($SPRTN^{+/+}$) and hypomorphic ($SPRTN^{f/-}$) mice revealed a 1.5 – 2-fold increase in dT-Tyr in the hypomorphic mice, providing direct evidence that Spartan plays a role in the repair of radical-induced DPCs.

Finally, we investigated the formation of formamidopyrimidine (FAPy) adducts after exposure to 3,4-epoxybutene, an epoxide metabolite of the known carcinogen 1,3-butadiene (**Chapter 5**). We successfully synthesized and structurally characterized a novel

BD-induced DNA adduct EB-FAPy-dG, and developed a sensitive isotope dilution tandem mass spectrometry assay for its detection *in vitro* and in cells. To our knowledge, this is the first report of a BD-induced FAPy adduct, and future studies will examine whether BD-induced FAPy adducts

In summary, during the course of this Thesis, we utilized mass spectrometry-based proteomics techniques to identify the proteins susceptible to PM- and ROS-induced DPC formation. After structurally characterizing the atomic connectivity of these adduces, we developed sensitive and accurate isotope dilution tandem mass spectrometry assays to perform absolute quantitation of PM- and ROS-induced DPC formation in cells and tissues. These assays were further utilized to begin investigating the repair mechanism of DPCs in cells and tissues, including providing direct evidence that the metalloprotease Spartan is involved in the repair of radical-induced DPCs. Finally, we detected EB-FAPy-dG formation *in vitro* and *in vivo*, the first evidence of 1,3-butadiene induced formamidopyrimidine formation.

TABLE OF CONTENTS

Lists of tables

List of schemes

Figures

List of Abbreviations

8-iso-PGF2α	8-iso-prostaglandin F2α
ACN	acetonitrile
AGT	O^6-alkylguanine DNA alkyltransferase
Ar	argon
BER	base excision repair
Boc-Cys-OH	Boc-L-cysteine
CP	cyclophosphamide
cTni	troponin-I
Cys-NOR-N7G	*N*-[2-[cysteinyl]ethyl]-N-[2-(guan-7-yl)ethyl]amine
DEB	1,2,3,4-diepoxybutane
DMEM	Dulbecco's modified eagle's medium
DPC	DNA-protein cross-link
DTT	dithiothreitol
dT-Tyr	2-amino-3-(4-hydroxy-3-((1-((2R,4R,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)methyl)phenyl)propanoic acid
EB	3,4-epoxy-1-butene
EB-FAPy-dG	N^6-(2-deoxy-D-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-*N*-(2-hydroxybut-3-en-1-yl) -formamidopyrimidine
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionization

EtOH	ethanol
FBS	fetal bovine serum
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GO	gene ontology
h	hour
H_2O_2	hydrogen peroxide
HCl	hydrochloric acid
HR	homologous recombination
kDa	kilodaltons
LAD	left anterior descending artery
MEC	mechlorethamine
MEF	mouse embryonic fibroblasts
min	minute
N7-EB-dG	N7-hydroxybuten-1-yl-dG
NEIL1	endonuclease 8-like 1
NER	nucleotide excision repair
NM	nornitrogen mustard
$^{\bullet}O_2^-$	superoxide
$^{\bullet}OH$	hydroxide anion
PARP	poly-(ADP-ribose) polymerase 1
PBS	phosphate-buffered saline
PDH-1Eβ	pyruvate dehydrogenase E1 beta subunit
PM	phosphoramide mustard

PMSF	phenylmethanesulfonyl fluoride
ROS	reactive oxygen species
SOD2	superoxide dismutase [Mn]
TFA	triflouroacetic acid
TMT	thermo mass tags
TIC	total ion current
UV	ultraviolet light
XP	xeroderma pigmentosum
XPA	xeroderma pigmentosum complementation group A
XPD	xeroderma pigmentosum complementation group D

1. Introduction

1.1. DNA damage and its biological consequences

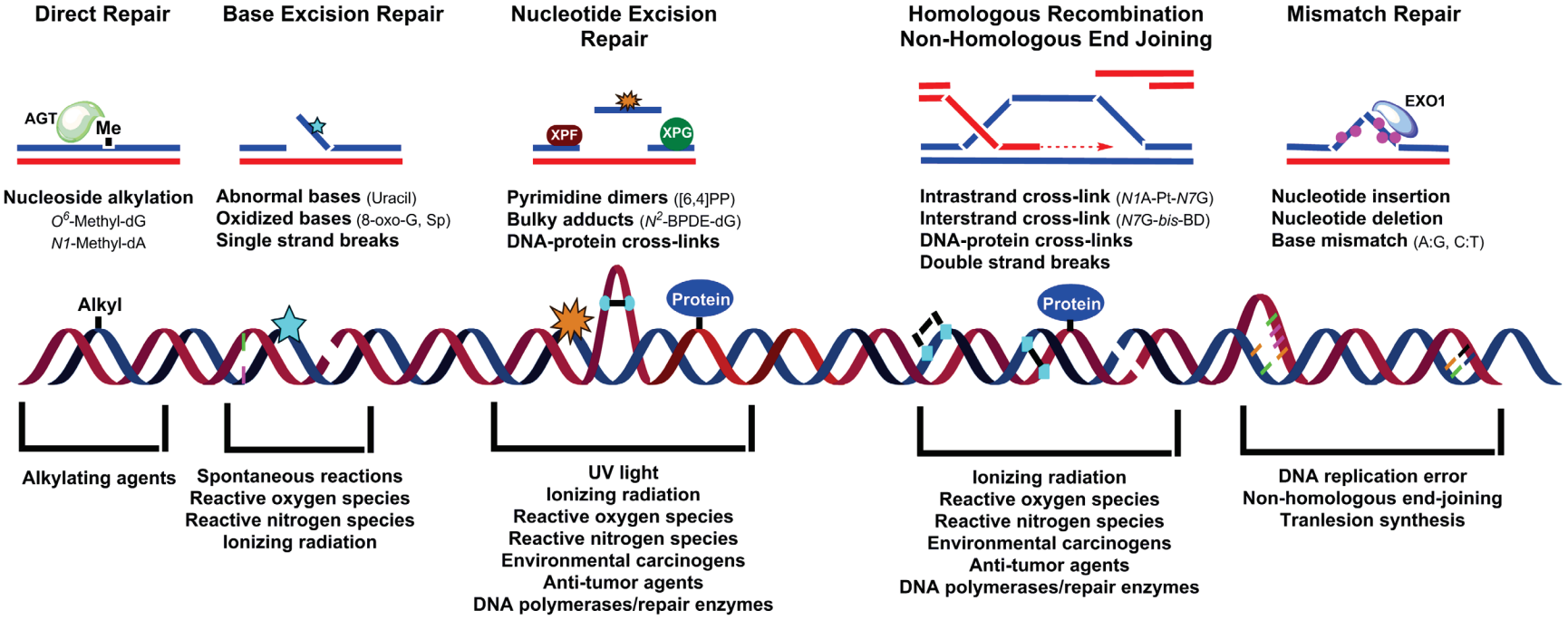
1.1.1. Structural modifications of DNA

Structural integrity of DNA is essential for the accurate expression and transmission of genetic information.[1, 2] However, DNA is constantly under threat of structural modification by endogenous and exogenous agents.[3] In addition to chemical reactions with oxidants and electrophiles, DNA is subject to spontaneous hydrolysis,[4] deamination of nucleobases,[4] and replication-induced errors. Modified nucleobases (DNA adducts) can be recognized and removed by cellular DNA repair enzymes that work to restore damaged DNA to its original state.[5-8] If not repaired, DNA adducts can induce errors during DNA replication or block DNA replication, resulting in mutations, senescence and apoptosis.[6] Therefore, the successful recognition and repair of DNA damage is crucial for cellular survival.

1.1.2. Sources of DNA lesions

DNA can be damaged by many physical and chemical agents including UV light,[9] ionizing radiation,[10] reactive oxygen species,[11] reactive metabolites of environmental carcinogens,[12, 13] common anti-cancer drugs,[3] and endogenously-produced reactive metabolites.[14] Exposure to these agents can lead to the formation of multiple types of DNA damage, including chemical modification to a single nucleoside (nucleobase monoadducts), DNA intra-strand and inter-strand cross-links (ICLs), DNA-protein cross-links (DPCs), single strand breaks (SSBs), and double strand breaks (DBSs, **Figure 1-1**).

Figure 1-1: DNA damage and repair pathways. Types of common DNA damage formed after exposure to endogenous and exogenous agents, and the known repair pathways involved in the removal of the damage.



Spontaneous deamination and hydrolysis of DNA are common events within human cells (100 – 400 events per genome per day).[4] Hydrolytic deamination of cytosine yields uracil.[15-18] This occurs via either attack of the neutral N^3-exocyclic amine by a hydroxide anion or attack of the protonated N^3-exocyclic amine by water.[15, 16, 19] Similarly, the N^6-exocyclic amine of adenine and the N^2-exocyclic amine of guanine can undergo deamination to yield hypoxanthine and xanthine respectively, although only at 2 – 3% of the rate of cytosine deamination.[20]

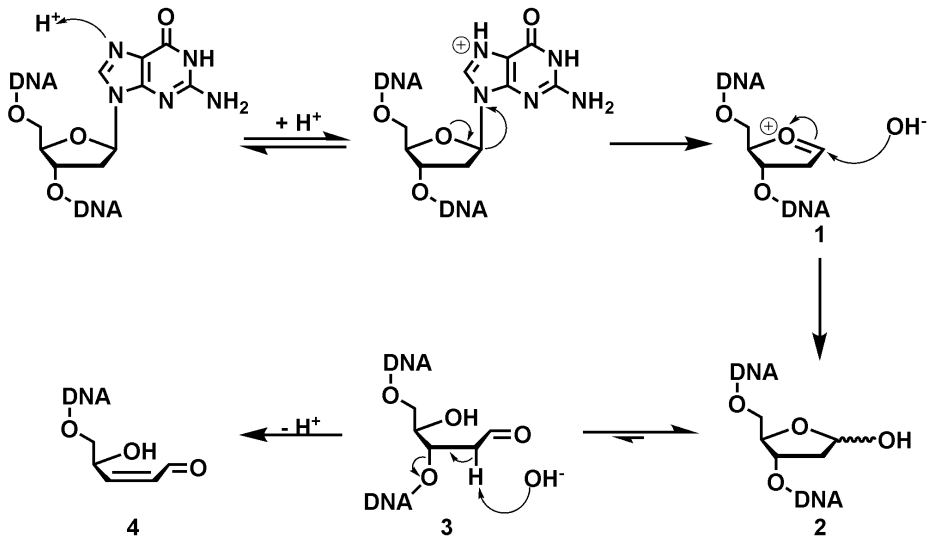
The glycosidic bonds of purine nucleosides of DNA can undergo acid-catalyzed hydrolysis, a process referred to as depurination.[21] Protonation of the N7-positions of purines stimulates the cleavage of the glycosidic bond, yielding the corresponding free bases and an oxycarbenium ion (**1**), which can undergo subsequent hydrolysis to yield an abasic site (AP site **2**, **Scheme 1-1A**).[21-23] It has been calculated that spontaneous depurination generates 10,000 AP sites per cell per day.[21] Conversely, the phosphodiester backbone of DNA is extremely stable under physiological conditions, with phosphodiester hydrolysis half-life estimated at 30,000,000 years.[24-26]

AP sites resulting from depurination are susceptible to further chemical reactions, yielding additional DNA lesions (**Scheme 1-1A**). Abstraction of an α-proton adjacent to the aldehyde group (**3**) leads to β-elimination of the 3'-phosphate group to generate DNA single strand breaks (**4**).[23, 27] The aldehyde group is also susceptible to further reaction with nucleophilic amino groups of other biomolecules (**Scheme 1-1B**). For example, the exocyclic N^2-amine of guanine residues on the opposite DNA strand can react with the aldehyde group of the lesion to form inter-strand cross-links (5'-dCAP, **5**) via Schiff base formation.[28] Similarly, reactive lysine residues in the catalytic sites of DNA polymerases

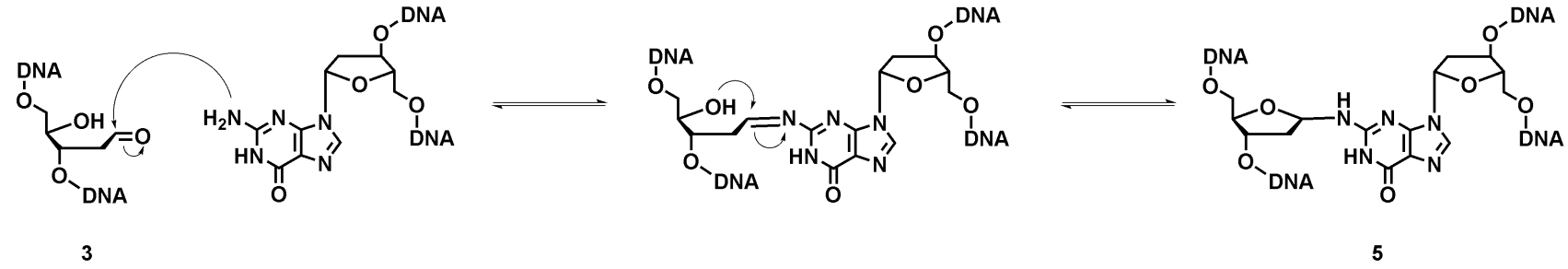
can react with AP sites of DNA to form covalent Schiff base DNA-protein cross-links.[29, 30]

Scheme 1-1: Acid catalyzed depurination of guanine and subsequent chemical reactions leading to DNA strand breaks and DNA-DNA cross-linking. (**A**) Mechanism of acid-catalyzed depurination and subsequent strand scission.[21-23] (**B**) Reaction of apurinic site **3** with a 2'-exocyclic amine of a neighboring guanine to form an interstrand DNA-DNA cross-link (**5**).[23, 27]

A.



B.



DNA can also be damaged as a result of exposure to reactive oxygen and nitrogen species (ROS and RNS).[3, 31, 32] ROS are produced via partial reduction of molecular oxygen. Examples of ROS includes superoxide anion ($^{\bullet}O_2^-$), hydrogen peroxide, (H_2O_2), and hydroxyl radical ($^{\bullet}OH$). RNS include nitric oxide ($^{\bullet}NO$), peroxynitrite ($ONOO^-$), nitrogen dioxide ($^{\bullet}NO_2$), and dinitrogen trioxide (N_2O_3). Exposure to ROS and RNS generates numerous DNA adducts including oxidized DNA bases, intra-strand and inter-strand cross-links, single and double strand breaks, and DNA-protein cross-links.[3, 31-33]

Over 100 different oxidatively damaged bases have been characterized *in vitro* (**Figure 1-2**).[31, 34] For example, hydroxyl radical can abstract hydrogens from the 2'-deoxyribose sugar backbone and DNA nucleobases, as well as add to the double bonds of DNA bases.[35] All five carbon positions of the 2'-deoxyribose sugar (dR) are susceptible to proton abstraction by $^{\bullet}OH$ to yield direct strand scission and electrophilic moieties.[36-38] Examples of dR-proton abstraction products include 3'-ene-lactone (C'1), DNA erythrose adducts (C'2), 3'-lactone (C'3), 4'-keto-1'-aldehyde (C'4), and DNA 5'-aldehyde (**Figure 1-3**).[38] Due to steric factors, the C'-5 and C'-4 positions are the most accessible to hydrogen abstraction.[39] The resulting electrophilic lesions can form covalent cross-links with nucleophilic sites on DNA and proteins, generating covalent DNA-DNA and DNA-protein cross-links.

Figure 1-2: ROS and RNS-induced DNA monoadducts, DNA tandem lesions, and DNA cross-links.

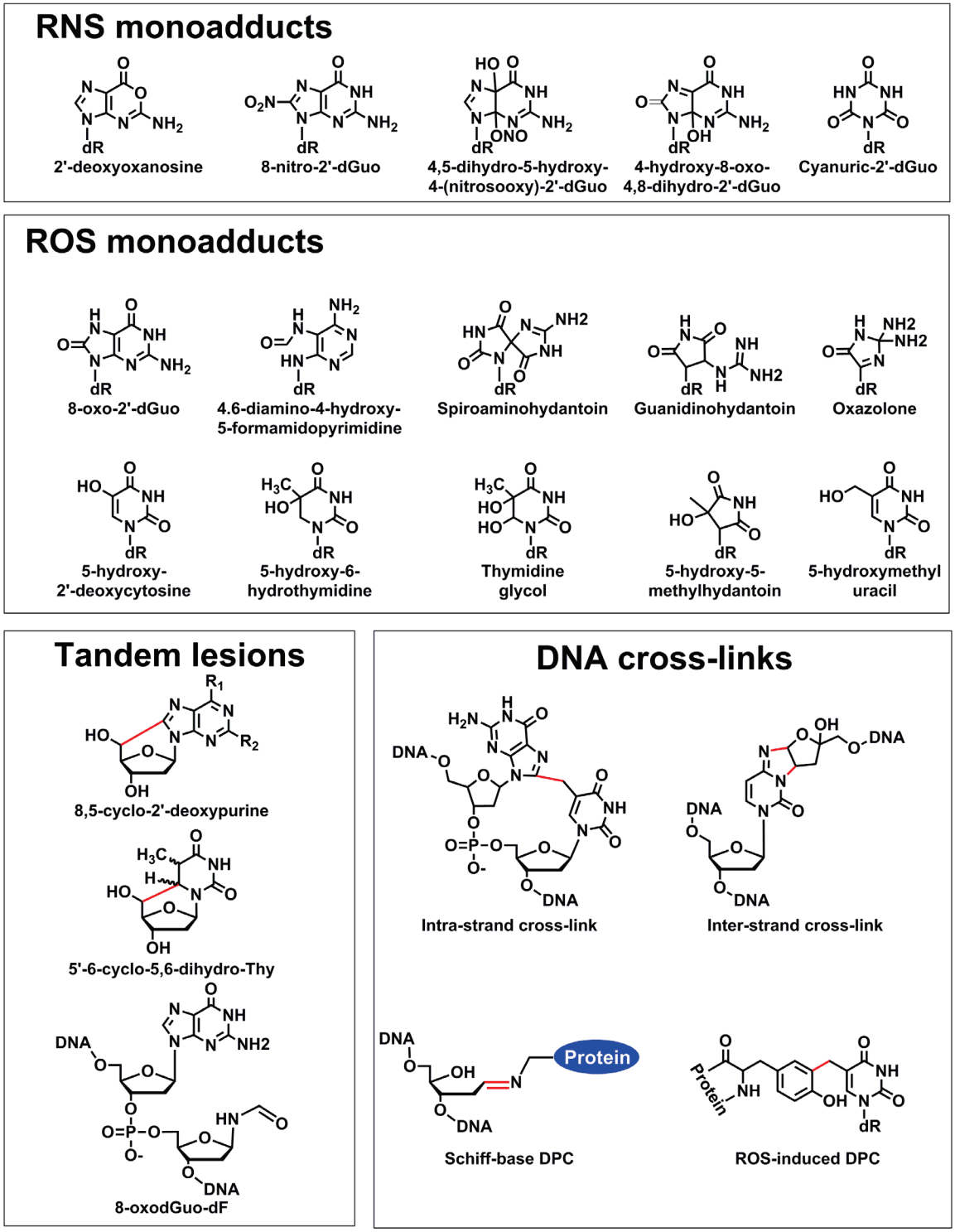
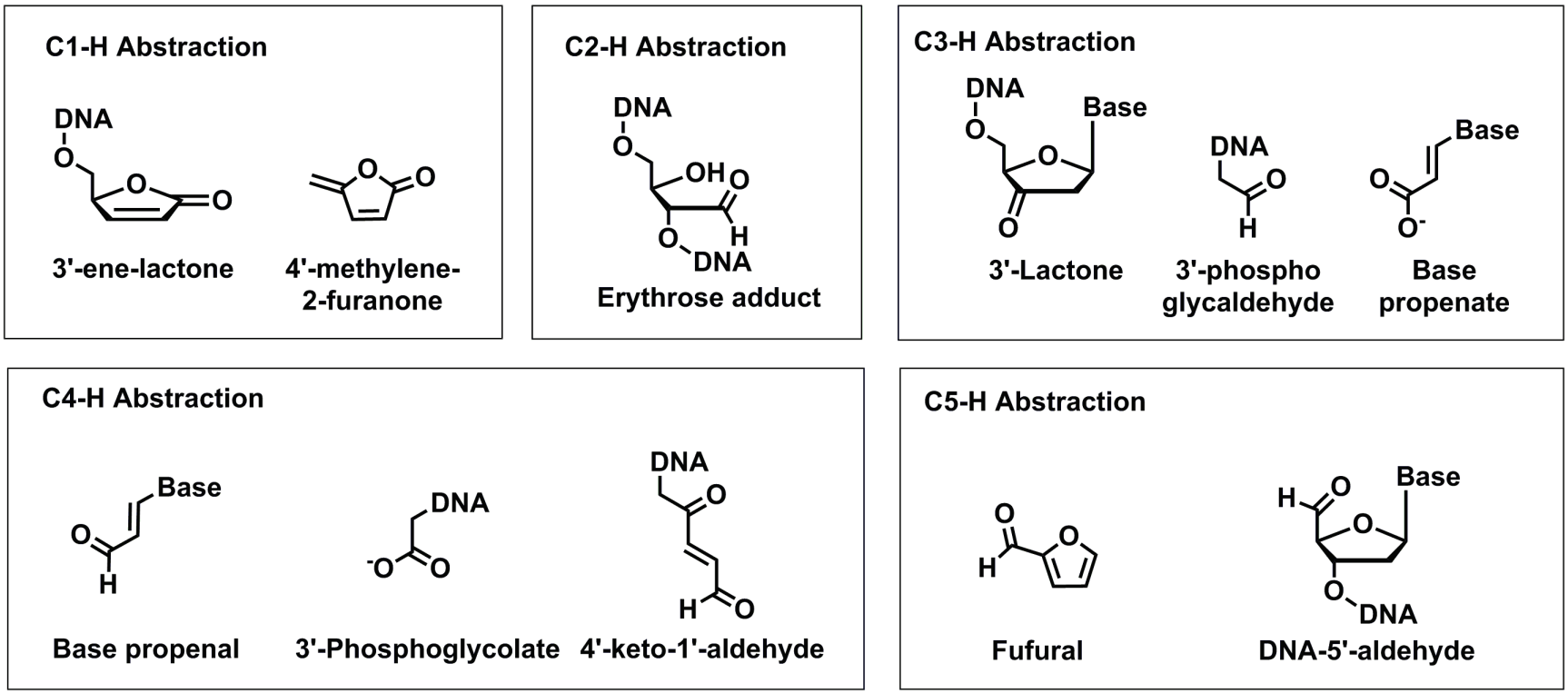


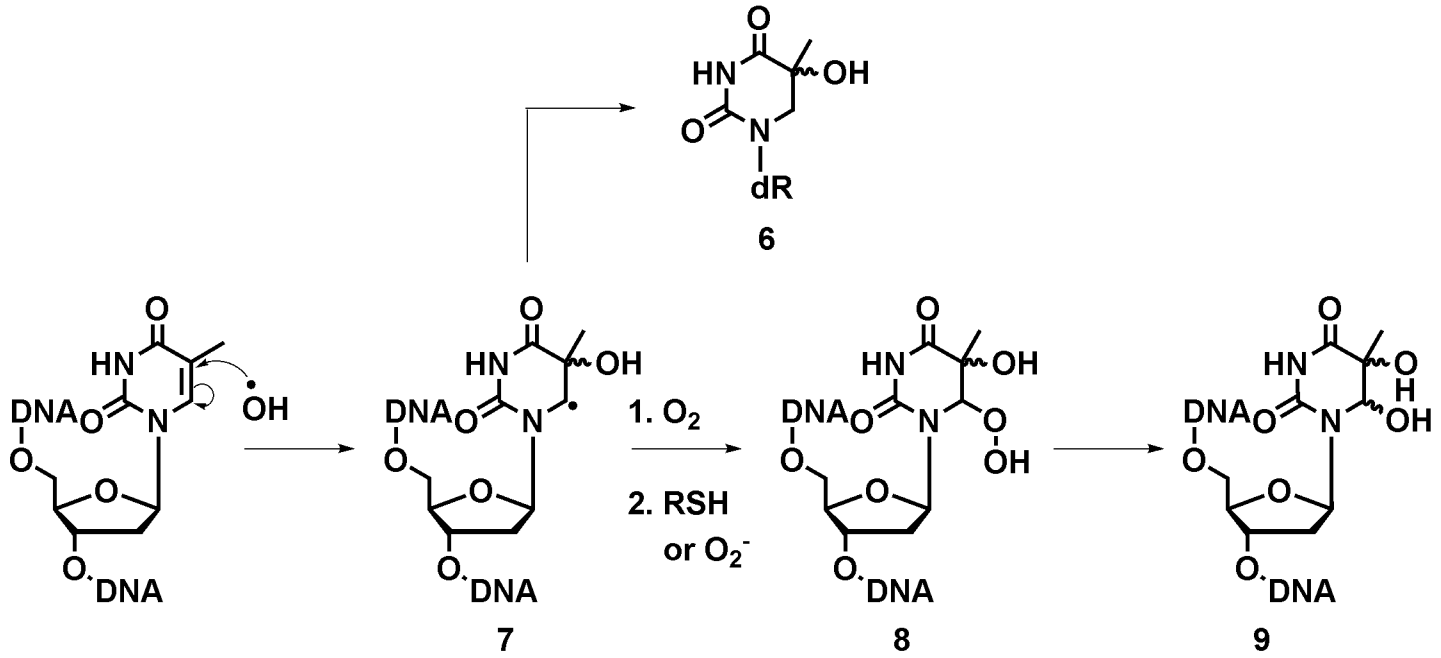
Figure 1-3: Potentially electrophilic DNA adducts formed after hydrogen abstraction from 2'-deoxyribose sugar of DNA.



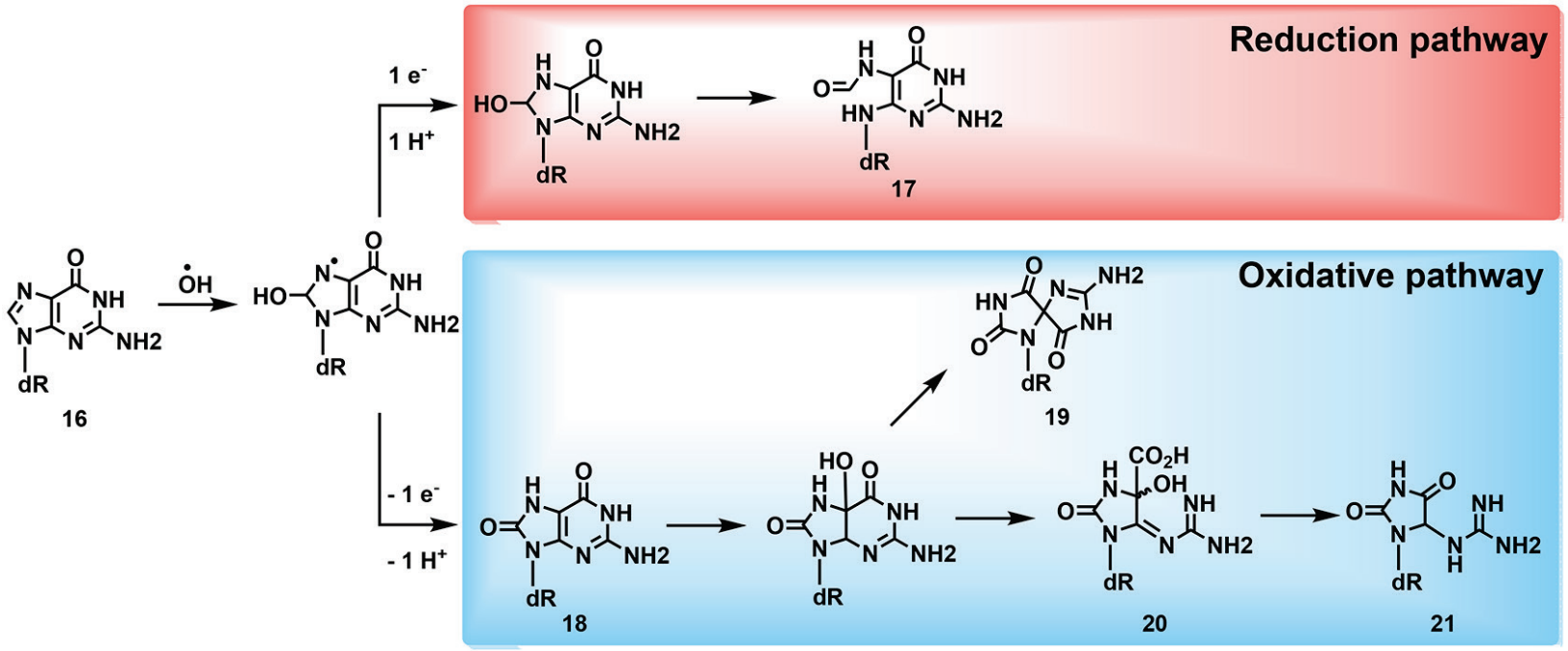
Some of the most prevalent oxidative base lesions are formed following hydroxyl radical addition across the π-bonds of bases.[40-42] Specifically, hydroxyl radicals can add across the 5,6-double bond of thymidine to yield a C6-thymidine radical (**6, Scheme 1-2**).[43] Radical **6** can be reduced to 5-hydroxy-6-hydro-thymidine (**7**) or react with molecular oxygen to yield a hydroperoxide (**8**), which will undergo subsequent reactions to yield the diastereomers of 5,6-dihydroxy-5,6-dihydrothymidine (**9**, Gly-Thy).[36, 40-42, 44, 45] Similarly, the C8-position of guanine can react with a hydroxyl radical to yield a nucleobase radical (**10, Scheme 1-3**) capable of undergoing one electron reduction or oxidation.[40, 46-48] Reduction of **10** results in imidazole ring cleavage to yield ring-opened formamidopyrimidine (**11**, FAPy) lesions. Oxidation yields the ubiquitous 8-oxo-7,8-dihydroguanosine (**12**, 8-oxo-dG),[49, 50] which is susceptible to further two electron oxidation and decomposition to spiroaminohydantioins (**13**), ureidoimidazolines (**14**), and guanidinohydantoins (**15**).[51-55]

ROS-induced DNA radicals can engage in secondary reactions with nearby bases in the DNA duplex, resulting in multiple damage sites and tandem lesions (**Figure 1-2**).[3, 31, 56] For example, a 5-(2'-deoxycytidinyl)methyl radical can react with the C8-position of an adjacent dG to yield G-T intrastrand cross-links (**Figure 1-2**).[49, 50] Other tandem lesions include 5,6-dihydrothymidine radicals generating an adjacent 2'-deoxyribonolactone,[57] 8-oxodGuo-dF,[58] and 8,5-cyclo-2'-deoxypurine cycloadducts.[59, 60]

Scheme 1-2: Mechanism of ROS-induced thymidine oxidation. Hydroxyl radical adds across the 5,6-double bond of thymidine to yield the reactive 5-hydroxy-6-hydro-thymidine radical (**6**), which be reduced to 5-hydroxy-6-hydro-thymidine (**7**) or further oxidized to 5,6-dihydroxy-5,6-dihydrothymidine (**9**, Gly-Thy).



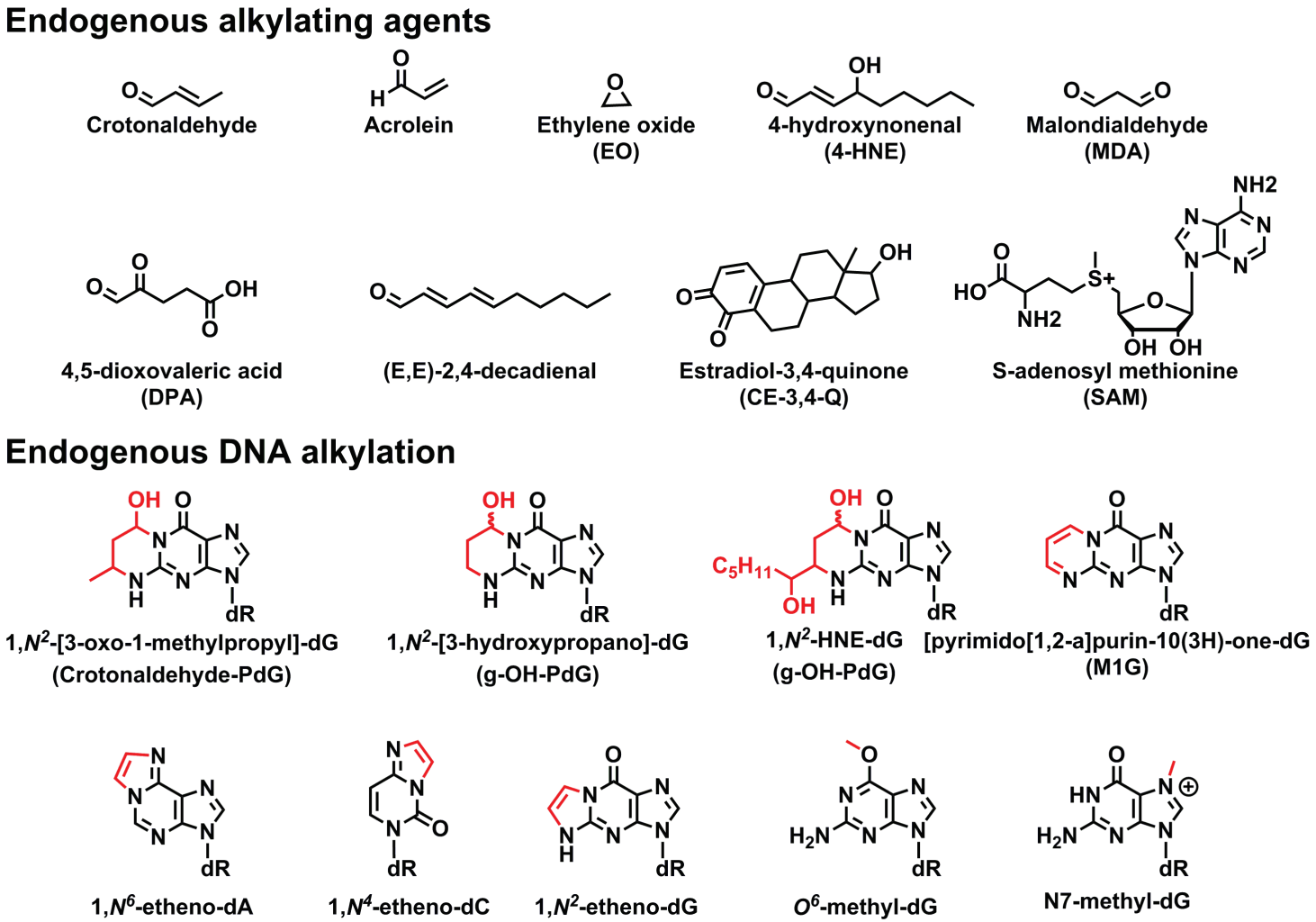
Scheme 1-3. Mechanism of ROS-induced purine imidazole ring opening. Attack of a hydroxyl radical at the 8-position of guanine generates an 8-hydroxy-guanine radical, which can undergo one-electron reduction to an unsubstituted formamidopyrimidine (FAPy) adduct, or one-electron oxidation to 8-oxo-dG (**12**) followed by further two electron oxidation and decomposition to spiroaminohydantioins (**13**), ureidoimidazolines (**14**), and guanidinohydantoins (**15**).



Chromosomal DNA can also become damaged after exposure to physical agents such as including ultraviolet light (UV) and ionizing radiation.[9, 61] UV irradiation can induce cyclobutane pyrimidine dimers (CPDs) and 6-4 pyrimidone photoproducts ([6-4]PPs) between adjacent pyrimidines.[62, 63] Together, these adducts account for nearly 75% and 25% of DNA damage after exposure to UV respectively. 6-4PPs can undergo further photoisomerization to Dewar valence isomers after exposure to wavelengths longer than 290 nm.[64] Ionizing radiation, such as X-rays and ɣ-rays can directly and indirectly damage DNA via ROS generation, forming the same oxidized DNA damage products described above (**Figure 1-2**).[31, 65]

Alkylating agents are a ubiquitous group of chemical agents capable of damaging DNA. These include industrial or household chemicals, environmental pollutants, known carcinogens, dietary chemicals, and chemotherapeutics. Endogenous alkylating agents such as methylglyoxal and malondialdehyde can be formed as byproducts of lipid peroxidation or other cellular processes (**Figure 1-4**).[14] Alkylating agents can react with ring nitrogens, exocyclic amines, and exocyclic oxygens of DNA bases to yield covalent DNA adducts (**Figure 1-4**).[5, 66] The specific type of DNA damage produced depends on many properties of the alkylation agent, including the number of reactive sites (monofuctional or bifunctional), the chemical reactivity of the agent (S_N1 vs S_N2 nucleophilic substitution), and the chemical structure of the group added.[6] The reactive moieties of alkylating agents can include alkyl halides, epoxides, aziridinium ions, aldehydes, and quinones.

Figure 1-4: Chemical structures of known endogenous DNA alkylating agents and the chemical structures of DNA alkylation products formed by these endogenous electrophiles.

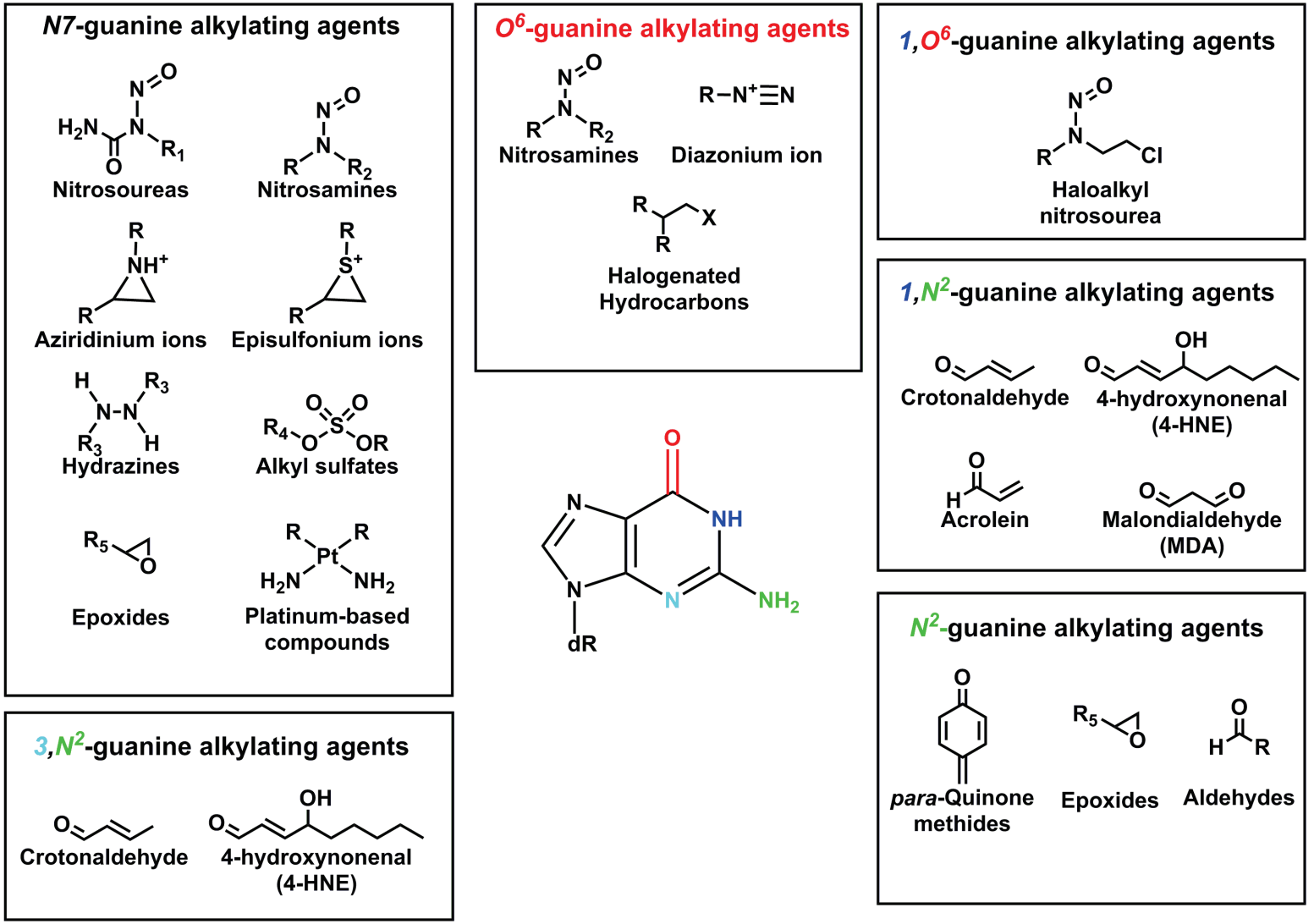


The most nucleophilic position on DNA is the N7-position of guanine (**Figure 1-5**).[67] Many small, monofunctional, and freely diffusible alkylating agents such as dimethyl sulfate predominately alkylate this position. Preferential alkylation of other sites has been justified by the hard-soft reactivity principle.[67-70] Hard alkylating agents, which are small, positively charged, and non-readily polarizable molecules such as diazonium ions, favor reaction with hard oxygen nucleophiles such as the C6-carbonyl position of purines (**Figure 1-5**). Softer alkylation agents, which are large, uncharged, and polarizable molecular such as chloroethylureas, preferentially alkylate ring nitrogens such as the *N*1 or *N*7 positions of purines (**Figure 1-5**).

DNA alkylation produces both hydrolytically stable and labile lesions, the latter are spontaneously released from DNA giving rise to abasic sites. Alkylation of the exocyclic amines of guanine (N^2-G), adenine (N^2-A), and cytosine (N^4-C) and the carbonyl oxygens of guanine (O^6-G) and thymidine (O^4-T) yields hydrolytically stable adducts.[71-74] Conversely, alkylation of the ring nitrogens of guanine (N7-G and N3-G), adenine (N7-A, N3-A, N1-A), and cytosine (N3-C) generates positively charged adducts that undergo depurination.[75-78] Alternatively, positively charged nucleobase adducts can undergo hydrolytic ring opening reactions to yield stable adducts. For example, N7-alkylated guanine and adenine are susceptible to hydroxyl anion attack at the C8-position of the purine heterocycle which initiates the cleavage of the imidazole ring to yield N^5-(alkyl)-formamidopyrimidine adducts.[78-80] N1-alkylated adenines can undergo transient pyrimidine ring opening via Dimroth rearrangement, where the alkyl group is transferred from the N1-position to the N^6-position of the heterocycle.[81, 82] Furthermore, *bis*-electrophiles can cross-link two nucleophilic sites of biomolecules, yielding DNA-DNA

and DNA-protein cross-links. Considering the vast number of DNA adducts generated in cells daily, dedicated repair mechanisms are required to prevent the accumulation and persistence of nucleobase damage.

Figure 1-5: Chemical structures of alkylating agents that react with 2'-deoxyguanosine to form monoadducts or cyclic adducts.



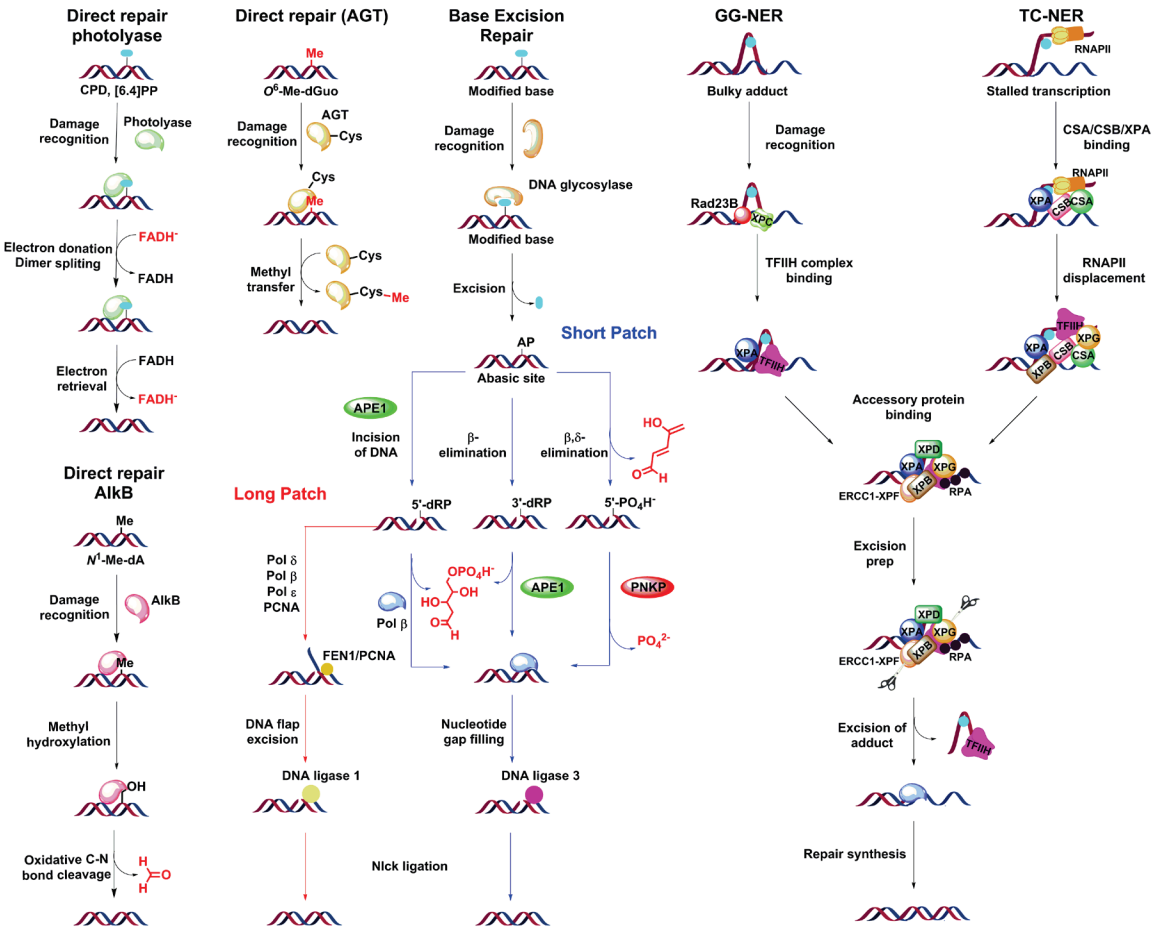
1.1.3. Repair of DNA damage

To ensure faithful replication and expression of genetic material, living cells have evolved multiple DNA damage response (DDR) pathways to recognize DNA damage and return DNA to its unaltered state. These pathways include direct repair, base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR), homologous recombination (HR), and non-homologous end-joining (NEHJ) (**Scheme 1-4 and 1-5**). The molecular size and shape, the degree of helix distortion, and the thermodynamic stability of each adduct influence their repair mechanisms.

Direct repair

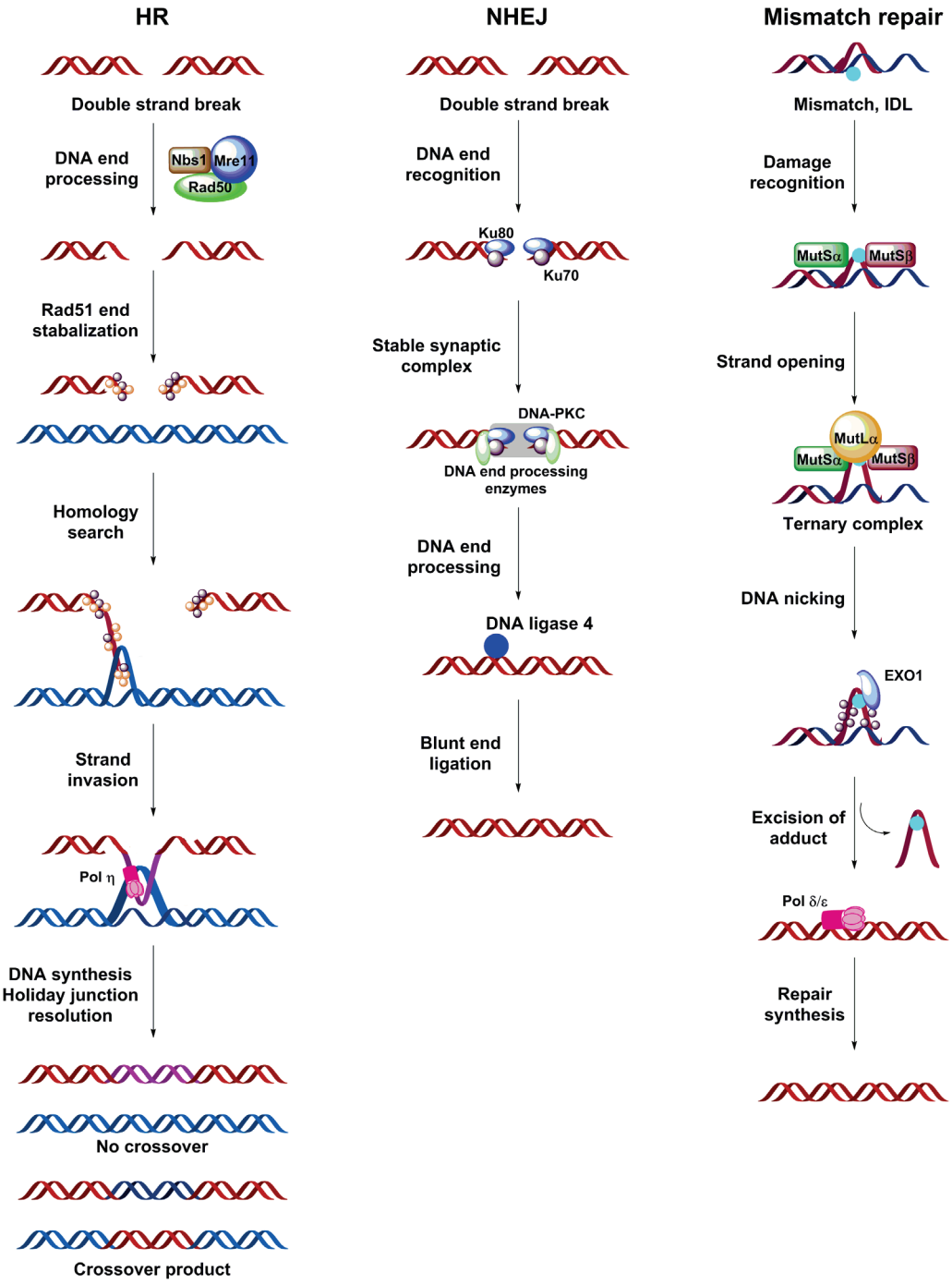
This type of repair involves direct reversal of damaged nucleobases to their native state by a single enzyme without the cleavage of the DNA phosphodiester backbone.[83] Three distinct direct repair enzyme classes have been identified: photolyases, O^6-alkylguanine-DNA alkyltransferase (AGT), and AlkB family dioxygenases. Bacterial photolyases reverse UV-induced photolesions such as CPPs and [6-4]PPs through a FADH⁻mediated electron transfer mechanism.[84, 85] AGT is a suicide DNA repair protein that repairs promutagenic O^6-alkylguanine and O^4-alkylthymine lesions.[86] Following protein binding to damaged DNA, the alkylated nucleotide is flipped out of the base stack into the AGT's active site, allowing for the alkyl group to be transferred to an activated side chain thiolate anion of Cys145, restoring native base.[87] AlkB family dioxygenases repair alkylated DNA damage through a Fe(II) and O_2-mediated oxidation and subsequent C-N bond cleavage.[5, 83, 88-92] N^1-methyladenine, N^3-methyl cytosine, N^1-methyl guanine, N^3-methyl thymine, 1,N^6-etheno adenine, 1,N^6-ethano adenine, and 3,N^4-ethenocytosine have been shown to be substrates for AlkB dioxygenases.[88, 93-98]

Scheme 1-4: Detailed mechanism(s) of DNA damage response pathways including; direct repair by bacterial photolyase, human AlkB, and human AGT; short patch and long patch base excision repair; global genomic and transcription coupled nucleotide excision repair.



AGT: O^6-alkylguanine-DNA alkyltransferase; APE1: apurinic/apyrimidinic endonuclease; PNKP: polynucleotide kinase 3'-phosphatase; FEN1: flap endonuclease 1: PCNA: proliferating cell nuclear antigen; RNAPII: RNA polymerase II holoenzyme

Scheme 1-5: Detailed mechanism(s) of DNA damage response pathways including: homologous recombination (**HR**) of a double strand break, non-homologous end-joining (**NEHJ**) of a double strand break, and mismatch repair of an insertion/deletion lesion.



Nbs1: nibrin; Ku70/80: x-ray repair cross-complementing proteins; DNA-PKc: DNA-dependent protein kinase, catalytic subunit; MutSα/β: DNA mismatch repair protein; EXO1: exonuclease 1.

Base excision repair

BER is the primary cellular DNA repair pathway responsible for removing deaminated bases, non-bulky alkylated nucleobases and oxidative lesions, single-strand breaks, and AP sites.[99-101] The nucleobase adducts are first recognized as BER substrates via a "base flipping" mechanism and excised from the duplex by DNA glycosylases.[99, 102-104] These enzymes are classified as either Type 1 (monofunctional) or Type 2 (bifunctional) glycosylases based on their catalytic mechanism and ability to execute AP lyase strand cleavage.[100, 105] Uracil DNA N-glycosylase (UDG), thymine DNA glycosylase (TDG), and methyl-CpG-binding domain4 (MBD4) are examples of Type 1 glycosylases, while MutY (MYH), 8-oxo-guanine glycosylase 1 (OGG1), and Endonuclease VII-like 1 (NEIL1) are examples of Type 2 glycosylases (**Table 1-1**).

Type 1 DNA glycosylases catalyze the cleavage of the glycosidic bond between the damaged base and the 2'-deoxyribose sugar, releasing the free base and generating an AP site.[17, 99] The phosphodiester bond at the 5'-position to the AP site can then be cleaved by AP endonuclease 1 (APE1) to generate a DNA strand break with a priming 3'-OH group and the non-conventional 5'-deoxyribose sugar. Type 2 glycosylases possess 3'-endonuclease activity and can catalyze both the cleavage of the glycosidic bond and incision of the phosphodiester bond 3' to the AP site.[99, 106] Type 2 glycosylases utilize an active site amine moiety to generate a covalent Schiff base between the enzyme and the AP site formed during the base excision step, keeping the substrate in position to allow for lyase activity.

The resulting abasic sites can be repaired by one of two mechanisms, referred to as short-patch BER (SP-BER) and long-patch BER (LP-BER). In SP-BER, DNA polymerase

β inserts a single nucleotide to fill the gap and cleaves the remaining 5'-deoxyribose phosphate from the DNA.[107-111] The DNA strand is then sealed by DNA ligase 3 (LIG3), completing the repair.[112] In LP-BER, DNA polymerase δ (Polδ), DNA polymerase ε (Polε), and proliferating cell nuclear antigen (PCNA) perform "strand displacement synthesis", adding 2 – 12 nucleotides and displacing the existing 5'-DNA end.[113] The displaced DNA is then cleaved by flap endonuclease 1 (FEN1), followed by LIG3-mediated ligation as above.[114, 115] The choice between SP-BER or LP-BER appears to be influenced by the concentration of available ATP after removal of 5'-deoxyribose, with higher ATP concentrations favoring SP-BER iniatiation.[116]

Recent studies have indicated that BER enzymes play a role in recognizing and removing ICLs formed by certain agents. HeLa cells deficient in NEIL1 enzyme were found to be hypersensitive to psoralen treatment.[117] Although psoralen mono-adducts have been identified as substrates for NEIL1 *in-vitro*,[117] McNeill *et al* showed NEIL1 is specifically recruited to ICL sites.[118] They treated HeLa cells with the psoralen derivative trioxsalen, which upon low laser micro-irradiation will stimulate ICL formation but will not induce ROS generation.[118] Furthermore, they treated HeLa cells with the psoralen-derivative angelicin, which can only form monoadducts upon the same laser micro-irradiation.[118] NEIL1 was recruited immediately to ICL sites, while angelicin treatment yielded no effect, showing NEIL1 recognized a specific feature of the psoralen-induced ICL.[118] Paradoxically, NEIL1 recruitment to ICL was also found to inhibit the excision of ICLs by XPC (NER pathway), suggesting the NEIL1's role in ICL repair is not fully elucidated.

Table 1-1: Mammalian BER enzymes and their substrates.

Enzyme	Subcellular localization	Type	Substrates
Uracil-DNA glycosylase 2 (UNG2)	Nucleus	I	Uracil: adenine mismatch Uracil: guanine mismatch 5-fluorouracil
Uracil-DNA glycosylase 1 (UNG1)	Mitochondria	I	Uracil: adenine mismatch Uracil: guanine mismatch 5-fluorouracil
Single-strand selective monofunctional uracil DNA glycosylase (SMUG1)	Nucleus	I	5-hydroxymethyl Uracil Uracil: guanine mismatch Uracil: adenine mismatch 5-fluorouracil 3,N^4-ethenocytosine
Thymine-DNA glycosylase (TDG)	Nucleus	I	Uracil: guanine mismatch Thymine: guanine mismatch 5-hydroxymethyl uracil 5-fluorouracil
Methyl-CpG-binding domain 4 (MBD4)	Nucleus	I	Uracil: guanine mismatch Thymine: guanine mismatch 5-hydroxymethyl uracil in CpG islands 5-fluorouracil, 3,N^4-ethenocytosine
3-methyl adenine-DNA glycosylase (MPG)	Nucleus	I	3-methyladenine, 7-methylguanine, 3-methylguanine, Hypoxanthine, 1,*N6*-ethenoadenine
8-Oxoguanine glycosylase (OGG1)	Nucleus	II	8-oxoguanine opposite cytosine FAPy-G opposite C
mutY DNA glycosylase (MUTY)	Nucleus	I	Adenine opposite 8-oxoguanine Adenine opposite cytosine Adenine opposite guanine
Endonuclease III-like protein 1 (NTHL1)	Nucleus	II	Thymine glycol, FAPy-G, 5-hydroxycytosine, 5-hydroxyuracil
Endonuclease 8-like protein 1 (NEIL1)	Nucleus	II	Thymine glycol, FAPy-G, FAPy-A, 8-oxoguanine, 5-hydroxyuracil, Dihydroxyuracil, Spiroaminohydantioins, Guanidinohydantoins
Endonuclease 8-like protein 2 (NEIL2)	Nucleus	II	Same as NEIL1
Endonuclease 8-like protein 3 (NEIL3)	Nucleus	II	FAPy-A, FAPy-G, Spiroaminohydantioins Guanidinohydantoins

Nucleotide Excision Repair

NER is the primary repair pathway used to remove bulky, helix distorting lesions including UV-induced cyclobutene pyrimidine dimers (CPDs) and (6,4)-pyrimidone photoproducts ([6,4]PPs), adducts of the environmental carcinogens such as benzo[*a*]pyrene and polyaromatic amines, and adducts of anti-tumor drugs such as cisplatin and phosphoramide mustard.[119-124] NER can be initiated by two subpathways, transcription-coupled NER (TC-NER) and global genome NER (GG-NER). TC-NER is responsible for rapid removal of distorting lesions from the transcribed strand of an active gene, and is initiated by the stalling of a RNA polymerase II.[125] Cockyane syndrome A (CSA), Cockyane syndrome B (CSB), and XPA binding protein 2 (XAB2) then facilitate the removal of the stalled polymerase to allow TC-NER to continue. GG-NER can occur anywhere in the genome, and is initiated by the GG-NER-specific XPC-RAD23B complex recognizing and binding a balky/distorting lesion. Specifically, the XPC protein binds the undamaged DNA strand and recognizes the two nucleotides opposite the lesion because they have increased ssDNA characteristics caused by thermodynamic destabilization of the adduct.[126] After damage recognition, both TC-NER and GG-NER share similar mechanisms and enzymes for completing the repair (**Table 1-2**).

The transcription factor IIH (TFIIH) complex, which consists of 10 subunits including XPB, cyclin-activated complex, and XPD, is recruited to the XPC-RAD23B complex.[127-129] The XPB protein will initially "open" the DNA strand, allowing the rest of the TFIIH complex to bind.[130] The XPD protein is believed to translocate across the damaged DNA strand, both unwinding the DNA and verifying the damage via helicase inhibition.[130-132] The loading of XPD allows the recruitment of RPA, XPA, and XPG to

complete the preincision complex.[129, 133] RPA binds to the undamaged strand, stabilizing the open "bubble" intermediate.[134, 135] XPA binds to the 5'-side of the bubble and interacts with TFIIH complex, RPA, XPC-RAD23B complex, DDP2, ERCC1-XPF complex, and PCNA, ensuring that these NER proteins are correctly positioned for the upcoming steps.[136, 137]

Once the endonucleases XPG and ERCC1-XPF are in position, they catalyze dual incision of the damaged strand, releasing a 24 – 32 nucleotide patch containing the nuclebase adduct and the TFIIH complex.[123, 138, 139] ERCC1-XPF performs the 5'-incision to generate a free 3'-hydroxyl group while XPG performs the 3'-incision.[138] The resulting gap is filled by either Pol δ/ε with a PCNA clamp or the TLS polymerase Pol ƙ, followed by DNA ligase 1 fixing the resulting nick.[140-142]

Table 1-2: Dynamic role(s) of nucleotide excision repair enzymes.

Enzyme	NER pathway	Step(s)	Activity	Protein binding
XPC	GG-NER	Damage recognition TFIIH loading	ssDNA binding	RAD23B, RAD23A, TFIIH, XPA, UV-DDB
RAD23A/B	GG-NER	Damage recognition TFIIH loading	Stabilize XPC Deliver XPC to UV-damage	XPC
CSA	TC-NER	Stalled fork recognition RNAPII removal	Ubiquitin ligase activity Recruit UV^{s}S-A-USP7 Protein binding	CSB, TFIIH (p44 subunit), XAB2, DDB2, UV^{s}S-A
CSB	TC-NER	Stalled fork recognition RNAPII removal	DNA binding/unwinding Protein binding	CSA, XPB, XPD, XPG, XAB2, TFIIH (p34 subunit), RNAPII
XAB2	TC-NER	Stalled fork recognition RNAPII removal	TBD	XPA, CSA, CSB, RNAPII
XPB	Both	TFIIH loading Pre-incision complex	ATPase DNA opening	TFIIH (p52)
XPD	Both	TFIIH loading Pre-incision complex	ATPase DNA helicase Damage verification	TFIIH (p44)
XPA	Both	Pre-incision complex Lesion excision	Protein binding Coordinate excision	TFIIH, XPC-RAD23B, DDB2, ERCC1-XPF, RPA, PCNA
XPG	Both	Lesion excision	Endonuclease	TFIIH, RPA
XPF	Both	Lesion excision	Endonuclease	TFIIH, ERCCI, XPA

Homologous Recombination

HR is a high fidelity, sequence-dependent repair process that preserves the genome by repairing complex DNA damage such as DNA gaps, double-strand breaks, and inter-strand DNA-DNA cross-links. HR also plays a prominent role in the faithful replication of the genome by supporting DNA replication and telomere maintenance. For all HR repair reactions, the key step is the homology search and 3'-DNA strand invasion by the Rad51-ssDNA presynaptic filament. This step ensures the new template DNA is in place to initiate the repair synthesis. A detailed explanation of DSB repair by HR is described below.

HR is initiated by the Mre11-Rad50-Nbsl (MRN) complex binding to DNA, followed by Mre11 enzymatically digesting DNA to creates 3'-DNA overhang regions at the site of the break. The resulting ssDNA region is quickly bound by RPA to prevent the formation of disruptive secondary hairpin structures. Rad51, with the aid of mediator proteins Rad54, breast cancer associated 2 protein (BRCA2), and Rad51 paralogs, will then bind to the ssDNA to form a nucleofilament (**Scheme 1-5**). Homology search and Rad51 recombinase-mediated strand invasion inserts the unaltered homologous strand from the sister chromatid across the damaged strand, an intermediate known as a D-loop. The homologous strand is then used as an error-free template for DNA synthesis by the TLS polymerase η. Ligation of the new oligo by DNA ligase 1 yields a four-stranded structure referred to as a Holiday junction. This junction is disassembled into either crossover or non-crossover products by Resolvase A or BLM DNA helicase with the type 1 topoisomerase TOPIII complex.

Non-homologous end joining

A competing repair pathway of DSBs is NHEJ, which directly ligates the two ends of a DSB.[143, 144] However, NHEJ does not consider sequence homology, and can lead to nucleotide additions and deletions. NHEJ is initiated by the Ku70/Ku80 heterodimer recognizing and binding the ends of the DSB, which both protects the ends from non-specific processing and ensures the recruitment of additional NHEJ proteins.[145, 146] Binding of DNA-PKCs to the Ku heterodimer forms a stable synaptic complex that holds the two broken ends together (add figure). This stabilized structure allows DNA end processing enzymes to remove end groups, fill gaps, and resect DNA ends to ensure the ends can be ligated together.[147, 148]

The removal of inhibiting end groups to yield ligatable ends is performed by polynucleotide kinase phosphate (PNKP), apraxtin and PNK-like factor (APLF), and the Ku heterodimer. PNKP is able to remove a phosphate from a 3'-hydroxyl group and to add phosphates from a 5'-hydroxyl group.[149] APLF catalyzes the removal of adenylate groups attached to a terminal 5'-phosphate.[150] The Ku heterodimer is believed to have 5'-dRP/AP lyase activity, particularly for AP sites within a 5'-overhang region of a DSB.[151] Resection of the DNA ends is performed by Artemis, APLF, and Werner (WRN), which create blunt ends that can be sealed together by DNA ligase 4.[152-155]

DNA mismatch repair

MMR is a highly conserved repair pathway that corrects DNA base-base mismatches that escape DNA polymerase proofreading, base-base mismatches generated by nucleobases deamination, oxidation, or alkylation, and insertion/deletion loops (IDLs) resulting from DNA polymerase slippage events.[156-158] The MMR pathway in *Esherichia*

coli has been extensively characterized biochemically and genetically, and has thus provided the framework for investigating the pathway in eukaryotes.[156, 159] In human cells, MMR is initiated by MutSα or MutSβ recognizing and binding to a DNA base-base mismatches or IDL in the unmethylated daughter strand.[158] MutSα preferentially recognizes mismatches and IDLs 1 or 2 nucleotides long, and MutSβ recognizes larger IDLs. Once bound to the mismatch, MutSα and MutSβ recruit MutLα and form a stable ternary complex that pulls the daughter strand out of the duplex.[160, 161] A portion of the single stranded DNA is bound by single-stranded binding-factor replication protein A (RPA),[162, 163] which allows exonuclease 1 (EXO1) to remove the DNA fragment containing the mismatch.[164] The resulting gap is filled by Pol δ and Pol ε, and the nick sealed by DNA ligase 1.[165]

1.1.4. Biological consequences of DNA damage

The accumulation and persistence of DNA adducts has been linked to the initiation, promotion, and development of diseases such as cancer,[166, 167] neurodegeneration,[168, 169] cardiovascular disease,[170, 171] and aging.[166] Unrepaired DNA adducts can contribute to these disease states by causing mutations in oncogenes or tumor suppressor genes, inhibiting DNA replication, and increasing chromosomal instability.

Mutations in genes associated with cell cycle checkpoints and DNA repair enzymes can lead to carcinogenesis. For example, mutations in the oncogenes *RAS* and *MYC* can lead to unregulated cell proliferation and eventually cancer.[172] Ataxia talengiectasia (AT) is a genetic disorder characterized by an increased risk of developing leukemia and lymphoma.[173] AT is caused by mutations to the *ATM* gene, which encodes the ATM

serine/threonine kinase that is activated by DSBs and key to the activation of DNA damage checkpoints.[173] Another genetic disorder caused by gene mutations is Xeroderma pigmentosum (XP). XP patients are unable to repair UV-induced DNA damage such as cyclobutane pyrimidine dimers (CPD) and pyrimidine-(6-4)-pyrimidone photoproducts ([6,4]pps) by NER, and are thus susceptible to sun-induced pigmentation abnormalities and an increased risk in skin cancers.[174]

Deficiencies in DNA repair pathways are also associated with neurodegeneration, premature aging, and carcinogenesis. For example, Cockyane's syndrome is caused by mutations to the *CSA* and *CSB* genes, and trichothiodystrophy is caused by mutations to the XPB and XPD genes.[175] Patients with both conditions are unable to execute TC-NER, leading to premature ageing and neurodegeneration.[175] Recently, germline mutations to the metalloprotease Spartan have been determined to be causative of Ruijs-Aalfs Syndrome, which is characterized by genome stability, premature aging, and early-onset hepatocellular carcinoma.[176] Spartan is currently being investigated as a crucial protease involved in DPC repair, making RJALS the first disease directly linked to the inability to repair DPCs. Considering the serious biological consequences of unrepaired DNA damage, is crucial to fully characterize DNA adducts and elucidate their repair mechanism(s).

1.2. 1,3-Butadiene-induced DNA damage

1.2.1. 1,3-Butadiene

1,3-butadiene (BD) is a colorless, volatile gas that is present in automobile emissions, urban and industrial air, and cigarette smoke.[172, 177-179] BD exposure in urban centers has been calculated between 1 – 10 ppb, while concentrations near petrochemical fueling centers have been calculated at over 100 ppb.[178-180] Comparatively, individual BD exposure from mainstream and side stream cigarette smoke has been estimated at 205 and 360 μg/cigarette respectively.[181, 182] BD is also an important raw material used in the production of synthetic rubber and plastics, especially tires used in the automobile industry.[178, 179]

Exposure to BD has been linked to many adverse health effects. Epidemiological studies have found that inhalation of BD is linked to an increased risk of leukemia, lymphatic and hematopoietic cancers, cardiovascular disease, and respiratory disorders.[183-190] Given these risks, the International Agency for Research on Cancer has classified BD as a known human carcinogen since 2008.[191] Despite the known risks of BD exposure, the molecular mechanism of BD-induced mutagenesis and carcinogenesis has not been fully elucidated.

1.2.2. Metabolism of 1,3-butadiene

The pathways of metabolic activation of BD to DNA-reactive species have been extensively studied over the years (**Scheme 1-6**). BD is first metabolized by the cytochrome P450 monooxygenases CYP2E1 and CYP2A6, which catalyze the epoxidation of the double bond to produce 3,4-epoxy-1-butene (EB).[192, 193] EB can either be hydrolyzed

by epoxide hydrolase (EH) to yield the unreactive 1-buten-3,4-diol (BED) or further oxidized to 1,2,3,4-diepoxybutane (DEB).[194, 195] The double bond of BED can be oxidized by CYP2E1 to yield 3,4-epoxybutan-1,2-diol (EBD), and one of the epoxide groups of DEB can be hydrolyzed by EH to yield EBD.[196, 197] All three possible stereoisomers of DEB (*S,S, R,R* and meso) are formed.[194] The epoxide metabolites of BD can be detoxified by glutathione conjugation/mercapturic acid pathway and excreted from the body in urine as a *N*-acetylcysteine conjugates (**Figure 1-6**).[198] If not detoxified, epoxide metabolites of BD can alkylate nucleophilic positions on biomolecules such as proteins and DNA.[199-203] The resulting BD-DNA adducts have varying mutagenic potentials depending on their molecular characteristics such as size, shape, hydrogen bonding potential, and thermodynamic stability.

Scheme 1-6: Metabolism of 1,3-butadiene and the formation of BD-DNA adducts

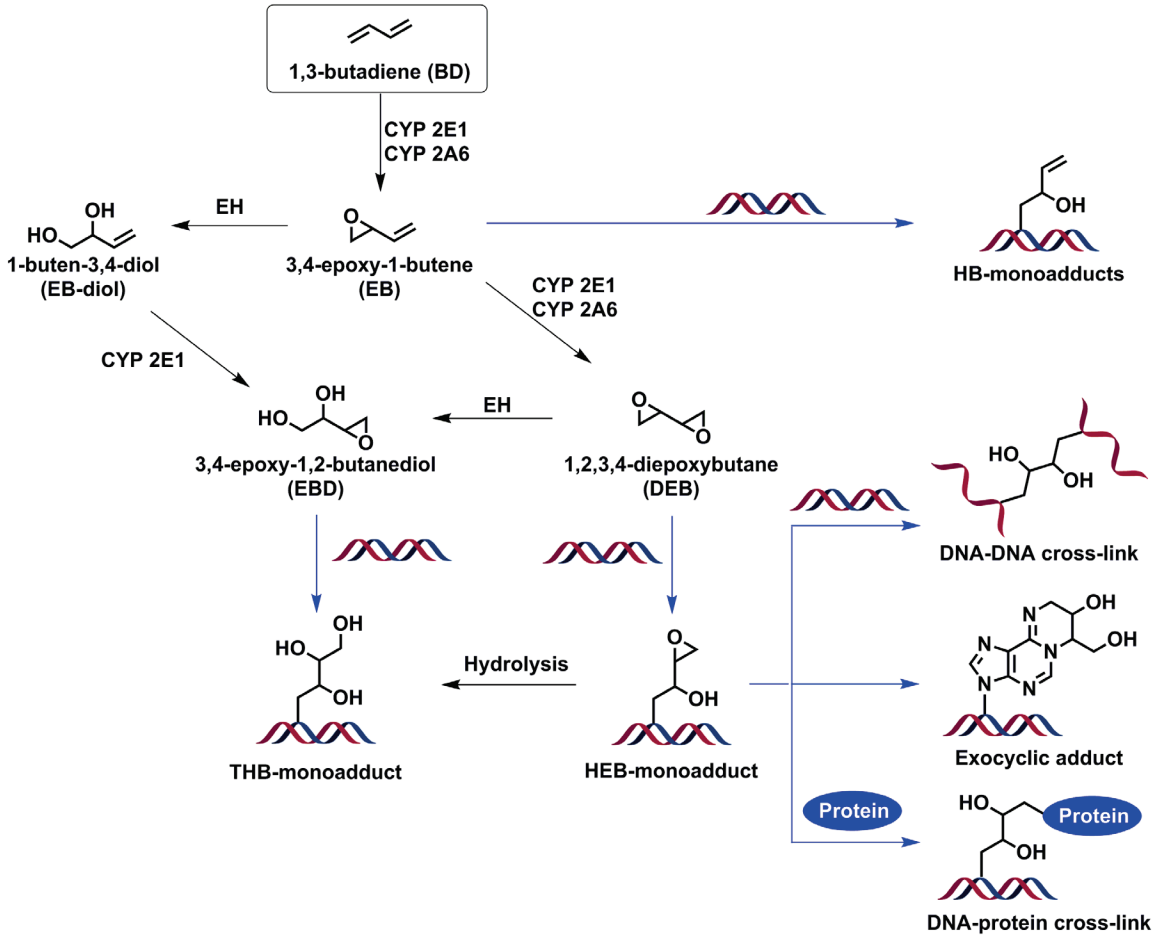
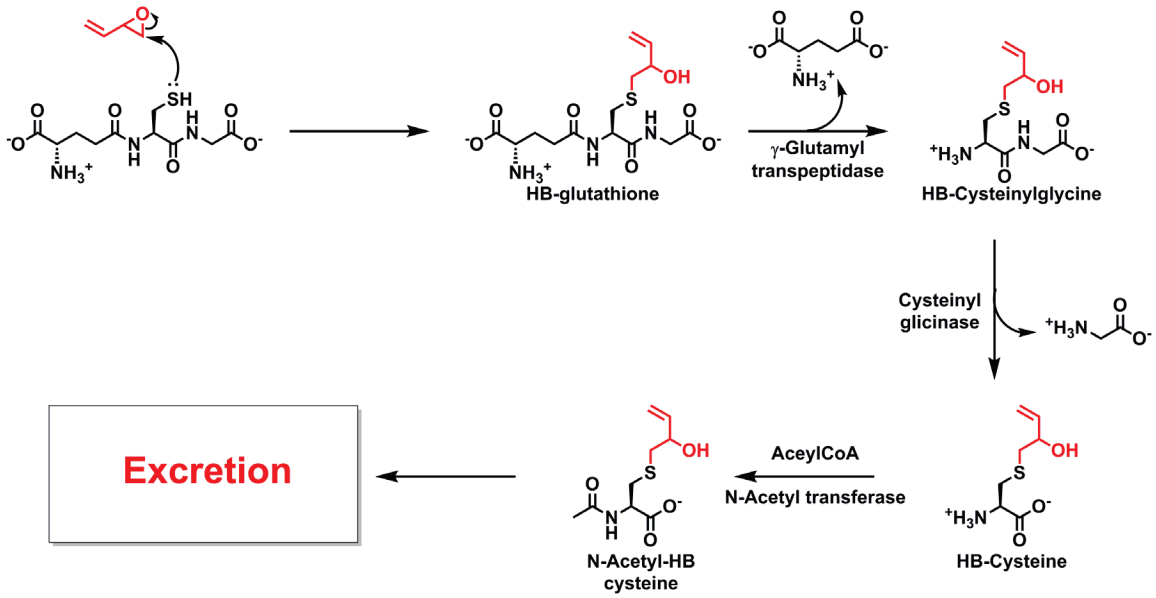


Figure 1-6: Detoxification of 1,3-butadiene metabolite epoxides through the mercapturic acid pathway.

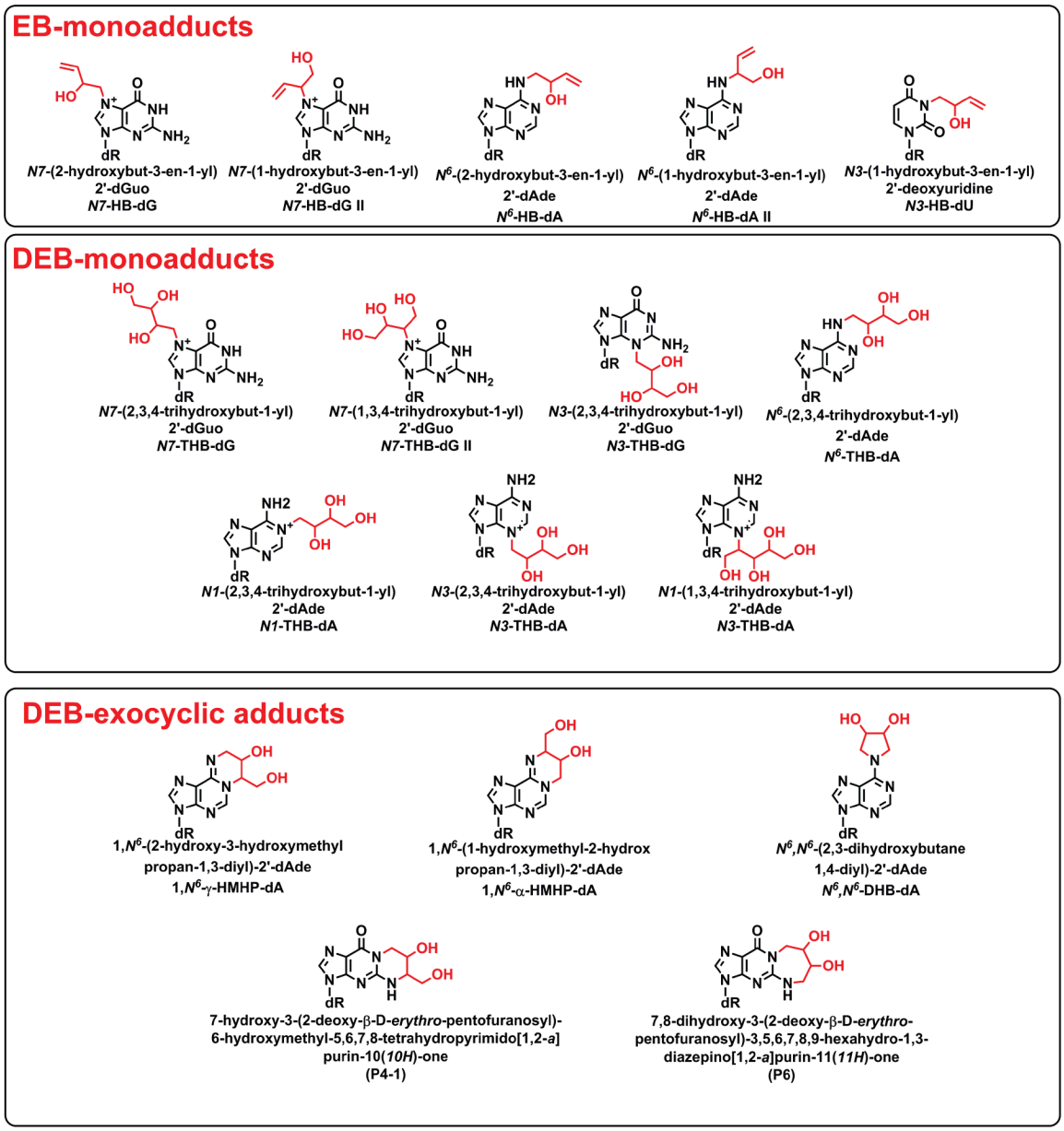


BD-DNA monoadducts

BD epoxides (EB, DEB, and EBD) are reactive electrophiles that can alkylate multiple positions of DNA including the *N7*-position of guanine, the *N3*-position of thymidine, and the *N1, N3,* and N^6-positions of adenine (**Figure 1-7**). Some examples of the resulting adducts include *N7*-(2-hydroxybut-3-en-1-yl)2'-dG (N7-HB-dG I),[200, 204] N^6-(2-hydroxybut-3-en-1-yl)2'-dA (N^6-HB-dA I),[200] *N3*-(2-hydroxybut-3-en-1-yl)2'-dU (N3-HB-dU),[205] *N7*-(2,3,4-trihydroxybut-1-yl)2'-dGuo (N7-THB-dG I),[206, 207] *N3*-(2,3,4-trihydroxybut-1-yl)2'-dGuo (N3-THB-dG I), *N1*-(2,3,4-trihydroxybut-1-yl)2'-dA (N1-THB-dA),[208] and N^6-(2,3,4-trihydroxybut-1-yl)2'-dA (N^6-THB-dA) (**Figure 1-7**).[208]

As described in section 1.2.2, DEB is a *bis*-electrophile that has two electrophilic epoxide moieties. DEB can react with two nucleophilic positions on the same nucleobase to form exocyclic lesions (**Figure 1-7**). The Tretyakova laboratory has identified exocyclic adenine lesions of DEB such as 1,N^6-(1-hydroxymethyl-2-hydroxypropan-1,3-diyl)-2'-Ade (1,N^6-α-HMHP-dA), 1,N^6-(2-hydroxymethyl-3-hydroxypropan-1,3-diyl)-2'-Ade (1,N^6-ɣ-HMHP-dA), and N^6,N^6-(2,3-dihydroxbutan-1.4-diyl)-2'-Ade (N^6,N^6-DHB-dA), with both 1,N^6-α-HMHP-dA and 1,N^6-ɣ-HMHP-dA observed in the tissues of BD-exposed rats (**Figure 1-7**).[209] The Elfarra laboratory reported the formation of exocylic guanine lesions 7-hydroxy-3-(2'-deoxy-β,D-*erythro*-pentofuranosyl)-6-hydroxymethyl-5,6,7,8-tetrahydropyrimido[1,2*a*]purin-10(1*H*)-one (H2) and 7.8-dihydroxy-3-(2'-deoxy-β,D-*erythro*-pentofuranosyl)-6-hydroxymethyl-3,5,6,7,8,9-hexahydro-1,3diazepino[1,2*a*] purin-10(11*H*)-one (H2'), although these lesions have not yet been observed *in vivo*.[210]

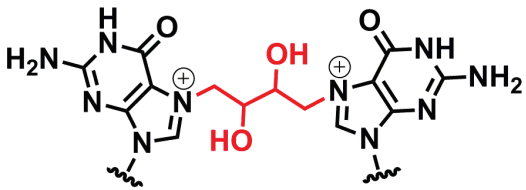
Figure 1-7: Structurally characterized nucleobase monoadducts and fused ring adducts induced by epoxide metabolites of 1,3-butadiene.



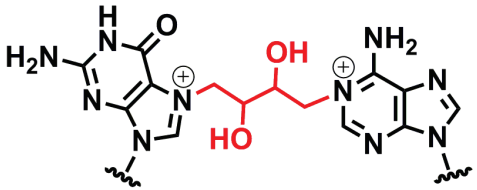
Intrastrand and interstrand DNA-DNA cross-links induced by DEB

Studies in cell culture have established that DEB is 50 -100-fold more mutagenic and genotoxic than other epoxide metabolites of BD (EB and EDB).[211, 212] This increased mutagenicity and genotoxicity is attributed to the ability of the diepoxide metabolite to form intrastrand and interstrand DNA-DNA cross-links. Among the three stereoisomers of DEB, (*S,S*)-DEB being the most mutagenic.[213, 214] The variability seen between the isomers is attributed to (*S,S*)-DEB preferentially forming the highly toxic 1,3-interstrand cross-link rather than 1,2-intrastrand cross-links.[215, 216] The most abundant DNA-DNA cross-link of DEB is 1,4-*bis*(guan-7-yl)butan-2,3-diol (*bis*-*N7G*-BD) (**Figure 1-8**).[203, 217] To a lesser extent, DEB also forms guanine-adenine cross-links 1-(aden-1-yl)-4-(guan-7-yl)butan-2,3-diol (*N1*A-*N7*G BD) and 1-(aden-3-yl)-4-(guan-7-yl)butan-2,3-diol (*N3*A-*N7*G BD) (**Figure 1-8**).[218]

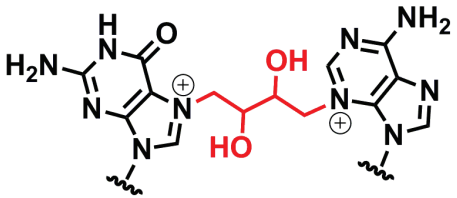
Figure 1-8: Chemical structures of DNA-DNA cross-links induced by the 1,2,3,4-diepoxybutane.



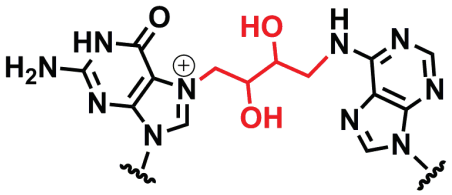
1,4-*bis*-(guan-7-yl)butan-2,3-diol
(*bis*-N7G-BD)



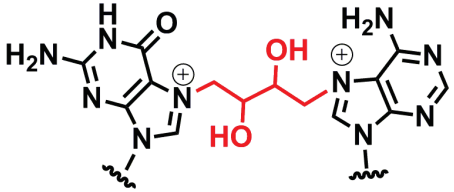
1-(aden-1-yl)-4-(guan-7-yl)butan-2,3-diol
(N1A-N7G-BD)



1-(aden-3-yl)-4-(guan-7-yl)butan-2,3-diol
(N3A-N7G-BD)



1-(aden-N^6-yl)-4-(guan-7-yl)butan-2,3-diol
(N^6A-N7G-BD)



1-(aden-7-yl)-4-(guan-7-yl)butan-2,3-diol
(N7A-N7G-BD)

1.3. N^6-(2-Deoxy-D-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-N^5-(alkyl)-formamidopyrimidine (N^5-alkyl-FAPy) adducts

1.3.1. Formation of FAPy adducts

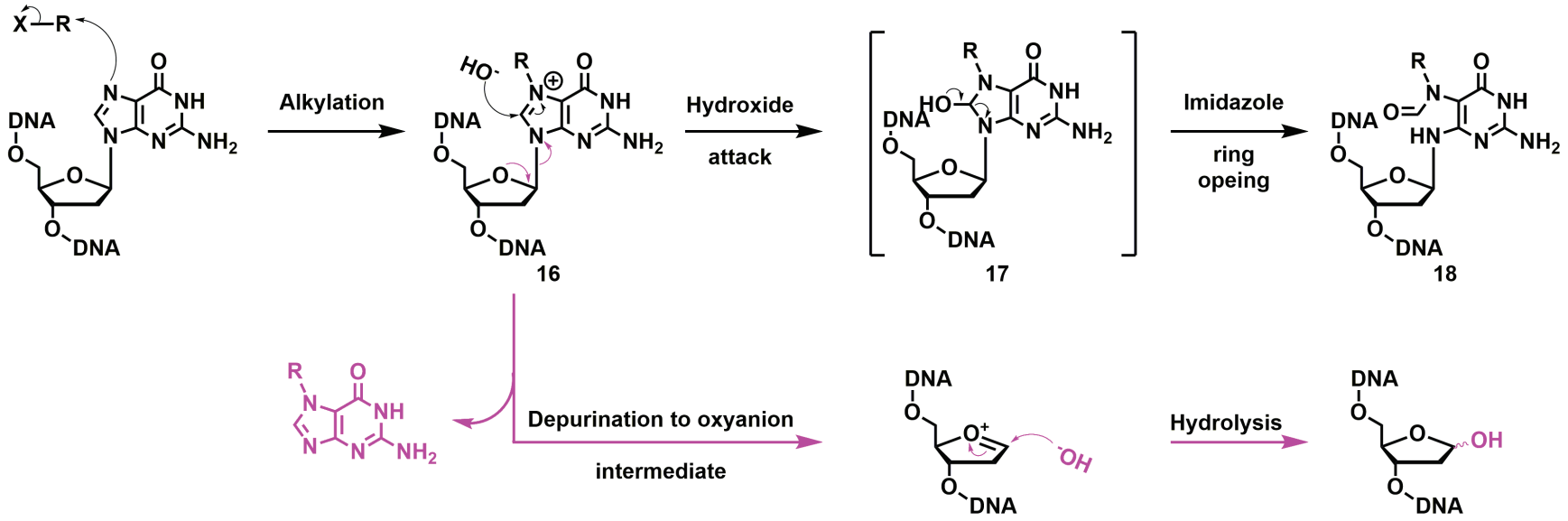
N^6-(2-Deoxy-D-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-formamidopyrimidine (FAPy) adducts can arise in DNA exposed to ROS (discussed in section 1.1.2) and after alkylation of the N7-position of purines. Alkylation of the N7-position of purine yields an unstable, positively charged N7-alkyl-dG cation adduct (**16**). The N7-alkyl-dG adduct can undergo two competing reactions: depurination to yield an AP site, or hydroxyl anion attack of the C8-position to yield an 8-hydroxy-N7-alkyl-dG intermediate (**17**), which immediately undergoes opening of the imidazole ring to yield the stable N^6-(2-Deoxy-D-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-N^5-(alkyl)-formamidopyrimidine (N^5-alkyl-FAPy) adducts (**18, Figure 1-9**).

N^5-alkyl-FAPy adducts have been observed after exposure to many electrophilic species. These include known carcinogens such as dimethyl sulfate,[219, 220] *N*-methylnitrosamines, *N,N*-dimethylnitrosamines,[221] environmental toxins such as aflatoxin B1(AFB1) epoxide[222, 223] and mitomycin C,[224] and antitumor and antibacterial compounds such as cyclophosphamide[225] and azinomycin.[226] All N^5-alkyl-FAPy adducts are characterized by significantly altered molecular structures and impaired hydrogen bonding potential as compared to the parent purine, leading to incorporation of incorrect DNA bases and/or DNA polymerase arrest during DNA replication.[227, 228]

As mentioned above, BD epoxides preferentially alkylate the N7-position of 2′-deoxyguanosine in DNA to form positively charged N7-guanine DNA adducts such as EB-Gua I, EB-Gua II and N7-THBG (see section 1.2.2).[200, 201, 206] Although no Fapy-dG

adducts of BD metabolites are reported in the literature, it is conceivable that EB, EDB, and DEB alkylation of purines could lead to the formation of N^5-alkyl-FAPy (e.g. EB-FAPy-dG and DEB-FAPy-dG), which could then contribute to the cytotoxicity and mutagenicity of BD.

Figure 1-9: Mechanism of N^5-alkyl-FAPy formation after alkylation of the N7-position of guanine. After alkylation of the N7-position of guanine, an unstable cation intermediate N7-alkyl-dG (**17**) is formed. The N7-alkyl-dG intermediate can undergo two competing reactions: depurination via oxyanion formation to release the alkylated guanine adduct and form an AP site (purple), or hydroxyl anion attack of the C8-position (**18**), immediately followed by the breaking of the imidazole ring to yield the stable (N^5-alkyl-FAPy) adducts (**19**).



1.3.2. Biological consequences of FAPy adducts

The formation of FAPy adducts dramatically changes the molecular shape and the hydrogen bonding characteristics of the parent nucleobase.[229-232] These structural modifications impair Watson-Crick base pairing preferences, leading to disturbances during DNA replication such as mutagenicity and genotoxicity.[227-230, 232-234] For example, N^5-alkyl-FAPy adducts have been shown to block DNA polymerases and induce G→T transversion mutations during DNA replication in SOS-induced *E. Coli*.[227, 228, 235] Similarly, N^5-alkyl-FAPy-dG adducts have been shown to inhibit DNA polymerases,[234, 236, 237] translesion synthesis (TLS) polymerases,[234] and induce G→T transversion in mammalian cells.[238] These increases in mutations may contribute to the mutagenic and toxic effects of known carcinogens, such as the epoxide metabolites of aflatoxin B1 and methylating agents. The mutagenicity and genotoxicity of several N^5-alkyl-FAPy adducts, most notably AFB1-FAPy-dG and Me-FAPy-dG, has been thoroughly investigated *in vitro* and *in vivo.*

AFB1-FAPy-dG

Aflatoxins are mycotoxins generated by certain molds including *Aspergillus flavus* and *Aspergillus parasiticus*, which grow in decaying vegetation, hay, and grains.[239] People can become exposed to AFB1 by consuming food stuffs that have been improperly stored, especially corn, rice, and wheat. There are four known aflatoxin isomers (aflatoxin B1, B2, G1, and G2, **Figure 1-10**), with aflatoxin B1 being the most hepatotoxic.[240] Acute AFB1 exposures can cause liver necrosis, while chronic exposures to AFB1 are linked to an increased risk of developing liver and gallbladder cancer.[241]

A combination of *in vitro* and *in vivo* studies have provided convincing evidence that AFB1-FAPy-dG is largely responsible for genotoxicity and mutagenicity of AFB1. Compared to N7-AFB1-dG, AFB1-FAPy-dG is significantly more stable in rat liver DNA.[242-244] Furthermore, AFB1-FAPy-dG adducts have been shown to stabilize double stranded DNA, presumable due to the intercalating ability of the AFB1 moiety.[229, 245] The major rotamer of AFB1-FAPy-dG, formed upon rotation around the C5-N^5 bond (**Figure 1-10**), has been reported to block *Salmonella typhimurium* DNA polymerase MucAB and *E. coli* polymerase UmuDC.[246] In cell culture studies, rat hepatocytes treated with AFB1 developed frequent G →T transversion mutations in the *Ki-Ras* oncogene, while human hepatocytes treated with AFB1 exhibited G →T transversion mutations in the *p53* oncogene.[247, 248] The latter results are consistent with G →T transversion mutations within codon 249 of the *p53* gene, which are observed in 50% of assayed hepatocellular carcinoma samples.[249, 250] Taken together, these results indicate that AFB1-FAPy-dG may be the major lesion responsible for mutagenicity of AFB1.

Me-FAPy-dG

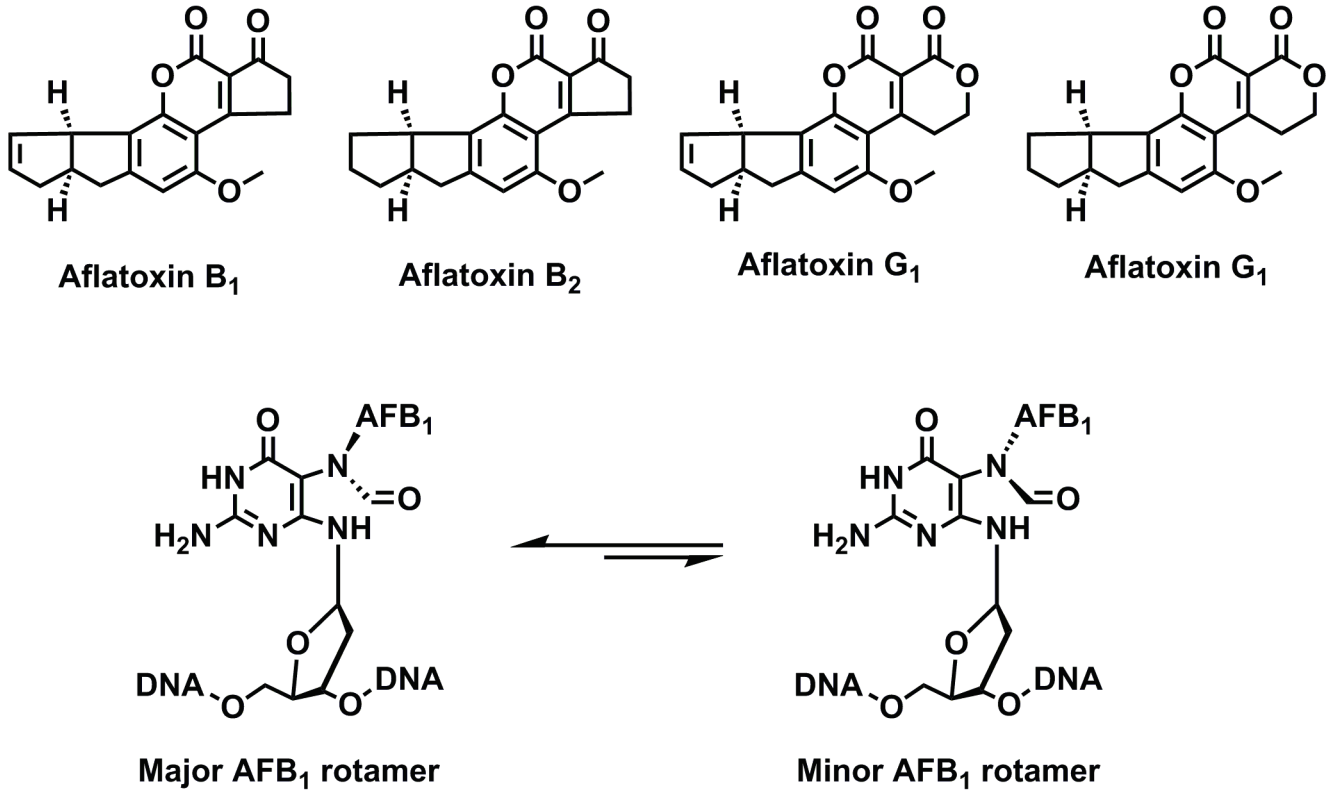
The N7-position of guanine in DNA is the primary alkylation site for small methylating agents such as dimethylsulfate and *N*-methylnitrosamines. The resulting N7-methyl-dG undergo base-catalyzed transformation to the corresponding Me-FAPy adducts.[79, 80] Me-FAPy adducts were first observed *in vitro* by Hanes *et al*, who methylated guanosine with diazomethane and structurally characterized the depurinated 2,4-diamino-6-hydroxy-N^5-methyl 5-formylamidopyrimidine (Me-FAPy) adduct. The first *in vivo* evidence for the formation of Me-FAPy adducts was provided by Beranek *et al*, who identified Me-FAPy in liver DNA of rats treated with *N,N*-dimethylnitrosamine and 1,2-

dimethylhydrazine, confirming that Me-FAPy-dG could form under physiological conditions.[251] Extensive NMR experiments have shown Me-FAPy can exist as a mixture of four rotational isomers due to the hindered rotation of the C5-N^5 and formyl-N^5 bonds.[251]

Recent *in vitro* studies have demonstrated the blocking and mutagenic potential of Me-FAPy-dG and Me-FAPy-dA adduct. *E. coli* DNA polymerase I exo and T4 DNA polymerase were completely inhibited one nucleotide before Me-FAPy-dG lesion.[236, 252] Ide *et al* investigated the bypass of a site specific methyl-FAPy-dG adduct by the Klenow fragment of *E. Coli* DNA polymerase 1 (KF) and observed that the FAPy lesion inhibited the extension step of lesion bypass.[237] Similarly, high fidelity eukaryotic DNA polymerases α, β, and human DNA polymerase δ were blocked at the insertion or extension step.[234] The TLS DNA polymerases η, κ, and Rev1/Polζ were able to efficiently bypass the lesion, although Pol η and Pol κ misinserted dT, dG, and dA opposite Me-FAPy-dG.[234]

Mutagenesis studies with M13mp18 phage plasmids in *E. coli* revealed that Me-FAPy-dA was 2-fold more mutagenic compared to Me-FAPy-dG following SOS induction.[227, 253] Specifically, M13mp18 phage DNA containing Me-FAPy-dG did not induce an increase in G → C transition mutations following SOS-induction. Conversely, the same phage DNA containing Me-FAPy-dA induced a two-fold increase in A → G transition mutations.[227, 253] These results suggest that methyl-FAPy-dA can mispair with 2'-deoxycytidine, while methyl-FAPy-dG is primarily a lethal lesion.[227]

Figure 1-10: Observed isomers of AFB-FAPy-dG adducts. Chemical structures of characterized aflatoxin metabolites (Aflatoxin B_1, AFB_1; Aflatoxin B_2, AFB_2; Aflatoxin G_1, AFG_1; Aflatoxin G_2, AFG_2) and the proposed major/minor rotamers of AFB_1-FAPy-dG adducts.



1.3.3. Repair of N^5-alkyl-FAPy adducts

Despite the growing evidence that N^5-alkyl-FAPy adducts are mutagenic, there is a relatively small amount of information regarding their cellular repair mechanisms. Several BER enzymes have been shown to remove N^5-alkyl-FAPy adducts. *E. coli* formamidopyrimidine DNA glycosylase (Fpg) efficiently removes unsubstituted FAPy-dG, unsubstituted FAPy-dA, and Me-Fapy-dG from DNA.[254, 255] The size of the N^5 substituent strongly influences this efficiency of Alk-FAPy repair by Fpg, as ethyl-FAPy-dG is removed nearly seven-fold slower than Me-FAPy-G.[256] 8-oxo-guanine DNA glycosylase (OGG1) also recognizes and repairs Me-FAPy-dG and unsubstituted FAPy-dG, as well as the carbocyclic analogue of Benzoyl-FAPy-dG.[257, 258] The only human BER enzyme known to remove unsubstituted FAPy-dA is NEIL 1, which can also remove Me-FAPy-dG.[259, 260]

NER has been identified as the repair pathway utilized to remove bulky AFB1-FAPy-dG adducts.[261] Plasmids containing AFB1-FAPy-dG adducts were transfected into wild type and repair deficient strains of *E. coli,* and host cell reactivation (HCR) assays were used to monitor lesion repair.[261] *E. coli* strains deficient in the NER protein uvrA were unable to repair the damage, while adduct repair in strains deficient in the BER protein mutM was less hindered.[261] Considering that the majority of published investigations have focused on unsubstituted FAPy-dG and Me-FAPy-dG, repair assays should be performed with additional N^5-alkyl-FAPy adducts to determine substrate specificity of key BER/NER enzymes, to discover how these adducts are ultimately repaired, and to help elucidate how these adducts could contribute to the mutagenicity of the parent compound.

1.4. DNA-protein cross-links

DNA-protein cross-links (DPCs) are ubiquitous, super-bulky DNA lesions that form when proteins become irreversibly trapped on chromosomal DNA.[262, 263] Due to their bulky nature, DPCs interfere with key biological processes such as DNA replication,[264-267] gene expression,[268] and chromatin packaging. DPC formation has been proposed to play a role in ageing,[269, 270] cancer,[271, 272] heart disease,[269, 270] and in neurodegenerative disorders.[273] Despite significant efforts to elucidate the biological effects of DPCs, their contribution to cytotoxicity and mutagenesis of *bis*-electrophiles is not fully elucidated. Furthermore, different cross-linking agents induce DPCs through distinct mechanisms, leading to distinct atomic connectivity and different groups of protein constituents involved in cross-linking. Finally, the mechanism(s) of active DPC repair remains controversial and is still not fully understood. Therefore, comprehensive characterization of the protein constituents and understanding the repair of DPCs is crucial to fully understand the biological consequences of DPCs.

1.4.1. Formation of DPCs in cells

DPCs can be induced after exposure to various physical and chemical agents including ionizing radiation,[274, 275] UV light,[276-278] transition metal ions such as Ni and Cr,[279-282] halogenated hydrocarbons,[283, 284] environmental carcinogens,[285] and common anticancer drugs such as nitrogen mustards and platinum compounds (**Figure 1-11**).[286-289] Endogenous DPCs can also form as a result of topoisomerases,[290, 291] DNA polymerases,[29, 30] and DNA repair enzymes becoming trapped on DNA intermediates, exposure to reactive oxygen species (ROS) produced during cellular respiration and inflammation,[292-294] and

exposure to aldehyde byproducts of lipid peroxidation such as methylglyoxal[295] and malondialdehyde.[296]

One of the most extensively studied DPC inducing agents is formaldehyde (FA). FA is a ubiquitous environmental carcinogen present in automobile exhaust, cigarette smoke, pharmaceuticals, resins coating particle board and plywood, and vapors from these products.[297] FA is also produced endogenously as a product of metabolic processes such as methanol oxidation, methylamine deamination,[298] and histone demethylation.[299, 300] Numerous studies have established FA as mutagenic, carcinogenic, and teratogenic compound.[301, 302] Although a number of DNA monoadducts and DNA-DNA cross-links have been observed in cells exposed to FA.[303] DPCs are considered the most abundant FA-induced lesion.[304]

Increased levels of FA-mediated DPCs have been found in the peripheral blood lymphocytes from occupationally exposed physicians, laboratory technicians, and orderlies.[272, 285, 305] Shaham *et al* observed a linear correlation between the years of occupational exposure and the levels of FA-induced DPCs in white blood cells, indicating DPCs could serve as a biomarker for chronic FA exposure.[285] Furthermore, concentration dependent FA-induced DPC formation has been observed in the nasal mucosa and respiratory tract of laboratory rats and rhesus monkeys exposed to FA by inhalation.[306, 307]

FA-mediated DNA-protein cross-linking takes place between lysine, histidine, tryptophan, and cysteine residues of proteins and dG, dA, and dC nucleosides of DNA.[308] Cross-linking initially generates Schiff base intermediates on proteins or DNA, which are converted to methylene cross-links upon reaction with another biomolecule (**Figure 1-11A**). *In vitro* cross-linking reactions between formaldehyde, 2'-deoxynucleosides, and

amino acids have identified dG-lysine and dG-cysteine cross-links as the most abundant adducts, although dG-lysine was unstable at ambient temperature.[308] At physiological conditions, the half-life of FA-mediated DPCs has been measured as 12.5 hours.[309] However, FA-induced DPCs are reversible upon heating via hydrolysis, a process exploited by biochemical assays such as chromatin immunoprecipitation (ChIP).[310-312]

Environmental and occupational exposure to the transition metal chromium (Cr) can induce DPC formation in cells and tissues.[280, 313] Cr-induced DPCs have been observed in the lymphocytes of chrome plating operators,[281] tannery workers,[281, 314] and stainless steel welders,[281, 313, 315, 316] as well as citizens exposed to Cr-contaminated water.[279, 281] Furthermore, Cr-induced DPCs have been observed in the erythrocytes of fish living in contaminated water.[317] Cr is present in metal alloys as the stable metallic Cr(0), but upon excessive heating or exposure to oxidative chemicals, Cr(0) will become oxidized to the water soluble Cr(III) and Cr(VI).[318] Due to its tetrahedral arrangement of oxygen groups, Cr(VI) structurally resembles sulfate and phosphate anions, allowing for its passive transport into cells via non-specific anion carriers.[319]

Once Cr(VI) enters the cell, it is quickly reduced to Cr(III) by ascorbic acid, glutathione, and cysteine residues.[320, 321] Cr(III) has six available coordination sites, easily allowing cross linking to the phosphate groups and N7 position of purines on DNA and the amino acid residues of proteins (**Figure 1-11B**).[322-325] Thiol containing cysteine and glutathione form the most abundant adducts, comprising 24 and 17% of all Cr(III)-DNA adducts in Chinese hamster ovary (CHO) cells treated with Cr(VI).[322] Additionally, histidine, glutamic acid, and ascorbate-Cr(III)-DNA adducts have been identified.[281, 322]

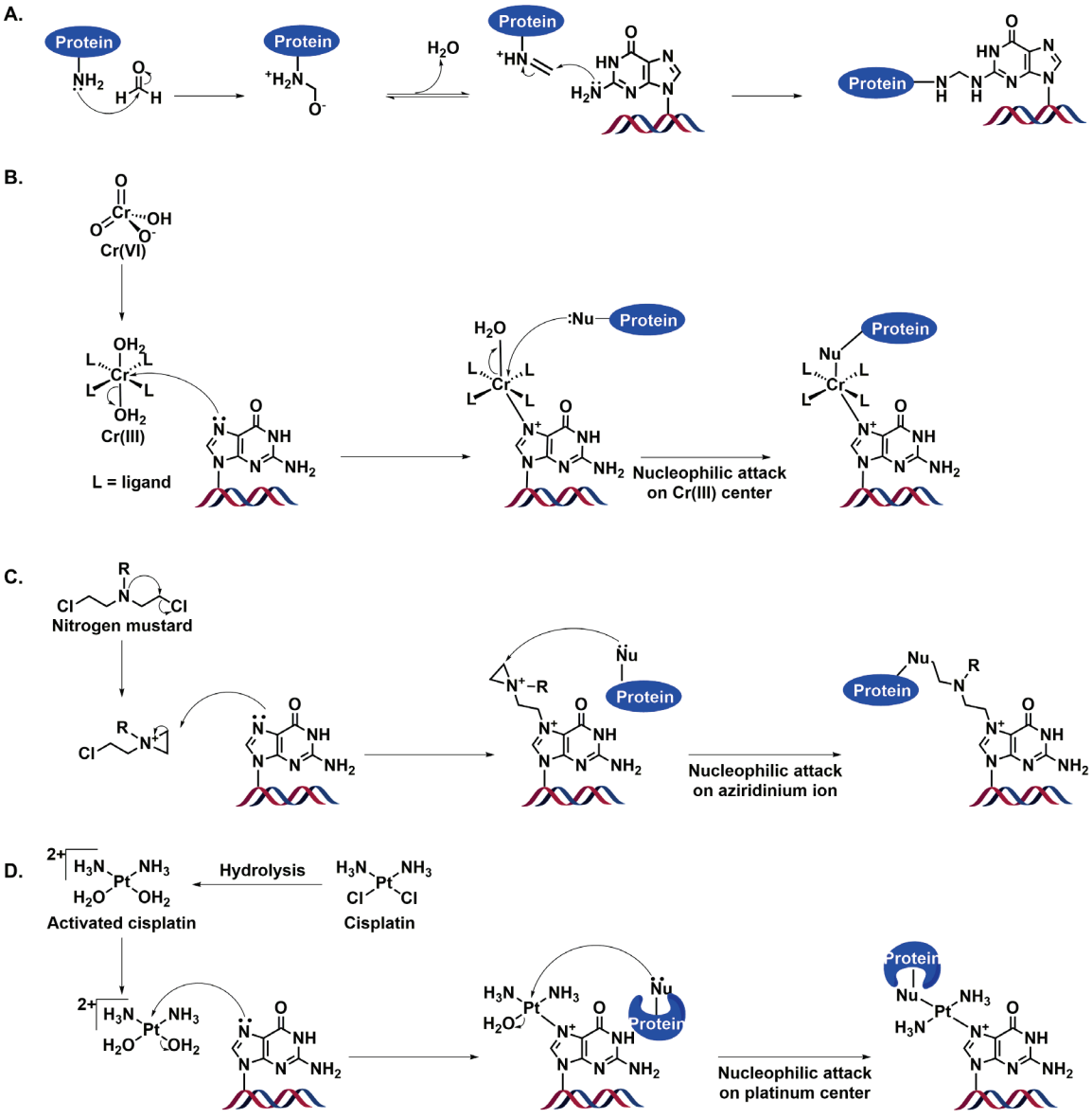
Nitrogen mustards (NM) are a class of anticancer drugs used as first-line chemotherapy treatments for lymphoma, leukemia, myeloma, breast and brain cancers, and in bone marrow transplantation.[326] Examples of clinically useful NMs include cyclophosphamide (CP), mechlorethamine, chlorambucil, melphalan, and ifosfamide.[327, 328] The first observation of NM-induced DPCs was made by Ewig *et al* [329], who detected DPCs in mouse leukemia cells (L1210) treated with nornitrogen mustard. Hansson *et al* [330] utilized the same methodology to observe nornitrogen mustard and melphalan-induced DPC formation in human melanoma (RPMI 8322) cells. Our laboratory employed the affinity capture technique mentioned above to identify 15 Chinese hamster ovary and 53 HeLa nuclear proteins that formed DPCs after exposure to mechlorethamine.[286, 287]

The antitumor effects of NMs are attributed to their ability to induce DNA-DNA and DNA-protein cross-links, which interfere with DNA replication, ultimately leading to cancer cell death. All NMs contain two N-(2-chloroethyl) groups, which can form reactive aziridinium ions capable of alkylating nucleophilic positions on DNA and proteins to form DPCs (**Figure 1-11C**). Using mass spectrometry analyses of model protein-DNA reactions, our laboratory has shown that the majority of NM-induced DPCs involve the N7 position of guanine and cysteine residues of proteins.[331]

Platinum agents such as *cis*-diamminedichloroplatinum(II) (cisplatin), carboplatin, and oxaliplatin are another class of commonly used cancer chemotherapeutics. Cisplatin was the first FDA-approved platinum anticancer drug, and is still widely used to treat bladder, head and neck, lung, ovarian, and testicular cancers.[332, 333] Although DNA-DNA intrastrand cross-linking is considered to be the major mechanism of activity of platinum agents, DPCs are likely to contribute to the observed therapeutic effects.

Chválová *et al* first observed platinum drug-induced DPC formation between histone H1 and the Klenow fragment of DNA polymerase with a synthetic DNA oligonucleotide.[334] For both proteins, Chválová *et al* observed much greater cross-linking with cisplatin compared to transplatin regardless of the length of the DNA oligo, incubation time, or concentration of drug.[334] Recently, our laboratory investigated cisplatin-induced DPC formation in human HT1080 cells and identified 256 distinct cross-linked proteins by mass spectrometry-based proteomics. [334, 335] Unlike other *bis*-electrophiles that are selective for cysteine side chains, cisplatin has been shown to form DPCs at various residues including arginine and lysine residues of proteins (**Figure 1-11D**).[334, 335]

Figure 1-11: Mechanisms of DPC formation by formaldehyde (**A**), chromium IV (**B**), phosphoramide mustard (**C**), and cisplatin (**D**).



1.4.2. Biological consequences of DPC formation

The contribution of DPCs to cytotoxicity and mutagenicity of *bis*-electrophiles was first demonstrated by Liu *et al.*[283] This group employed *E. Coli* cells overexpressing O^6-alkylguanine-DNA alkyltransferase (AGT). AGT is a DNA repair protein that protects the human genome from the damaging effects of promutagenic O^6-alkylguanine lesions.[86] Following protein binding to damaged DNA, the O^6-alkylguanine nucleotide is flipped out of the base stack into the AGT's active site, allowing for the O^6-alkyl group to be transferred to an activated side chain thiolate anion of Cys145.[87] AGT typically protects cells from toxicity associated with exposure to alkylating agents. However, when *E. Coli* overexpressing AGT were exposed to the bifunctional electrophiles 1,2-bromoethane (DBE) and diepoxybutane (DEB), toxic AGT-DNA cross-links were formed, leading to increased toxicity and mutations in cells overexpressing AGT.[283] These results suggested that covalent DPCs induced toxicity and genotoxicity in living cells.

One important limitation of the earlier studies is that cross-linking agents such as DBE and DEB form a range of DNA adducts such as DNA-DNA cross-links and monoadducts in addition to DPCs, making it difficult to directly attribute the observed biological effects to DPC formation. Our group developed a novel methodology to selectively induce DPCs into living cells.[336] In brief, recombinant human C145A AGT protein was treated with DEB to generate a DNA-reactive 2-hydroxy-3,4-epoxybutyl moiety on cysteine 150, and the resulting protein containing a DNA reactive epoxide group was purified by size exclusion. Electroporation of C145A monoepoxides into mammalian cell lines induced cell death and increased the frequency of mutations at the *hprt* gene (1.8 per 10^6 vs 1 per 10^6, $p<0.01$, Chi square). When C145A was replaced with a variant unable

to accumulate in the nucleus (K125L), lower toxicity and mutations were observed. Importantly, toxicity and mutation correlated with the numbers of covalent DPCs as detected by LC-MS/MS.[336] These results provided the first direct evidence that DPC formation leads to cell death.

Due to their super-bulky nature, DPCs are expected to inhibit crucial biological processes such as DNA replication and transcription, interfering with the faithful copying and expression of genetic information. The size of the protein constituent of DPC and the cross-linking site on DNA largely determine whether they present a complete block or can be bypassed by DNA and RNA polymerases. Nakano *et al* reported that DPCs over 14.1 kDa in size located on the translocating strand completely blocked DNA unwinding by human helicases (mini chromosome maintenance Mcm467 subcomplex), while DPCs between 5 to 14.1 kDa severely blocked helicase progression (**Table 1-3**).[264]

Table 1-3: *In-vitro* determination of the size limits for different helicase enzymes to bypass DPCs.

			Helicase inhibition by DPC		
Protein	**MW (kDa)**	**DnaB**	**T7gp4**	**Mcm467**	**Tag**
Oxanine control	0	None	None	None	None
Adrenomedullin (22 – 52)*	3.6	None	Minor stalling	None	Minor stalling
Parathyroid hormone (1 – 22)*	3.7	None	Minor stalling	None	Minor stalling
Gastric inhibitory peptide	5.0	None	Minor stalling	Severe stalling	Minor stalling
Midkine (60 – 121)*	6.8	Severe stalling	Severe stalling	Severe stalling	Severe stalling
Histone H2A	14.1	Blocked	Severe stalling	Blocked	Severe stalling
T4 endonuclease V	16.0	Blocked	Blocked	Blocked	Blocked
High mobility group protein 1,2	24.9	Blocked	Blocked	Blocked	Blocked
NEIL1 DNA glycosylase	43.7	Blocked	Blocked	Blocked	Blocked

Our laboratory investigated the effects of DPCs and smaller DNA-peptide conjugates on DNA replication *in vitro* by human DNA polymerases.[266, 267] Hydrolytically stable model DPCs were site-specifically incorporated into DNA using reductive amination reactions between protein's lysine residue and 7-deaza-7-(2,3-dihydroxypropan-1-yl)-2'-deoxyguanosine in DNA[337] or copper catalyzed [2+3] cycloaddition between azide-containing proteins and 5-(octa-1,7-diynyl)-uracil in DNA.[265] In vitro DNA replication experiments have uncovered complete blockage of human TLS polymerases η, κ, β, and ι by full size DPC lesions.[266, 267] However, when the protein was replaced with a 10-mer peptide, the polymerases were able to bypass the adduct.[265, 266] Polymerase bypass of 10-mer peptides attached to C5 position of cytosine in DNA by hPol η and hPol κ was error prone, resulting in both base substitutions and deletions.[266] In contrast, the same lesions conjugated to the N7 position of DNA were not mispairing.[267] More recently, we investigated the effects of bulky DPC lesions on transcription (Ji *et al,* unpublished data). We found that bacterial RNA polymerases were blocked by protein lesions conjugated to the N7 position of guanine. Lesions placed in the non-transcribed strand of DNA had a smaller effect on transcription (Ji *et al,* unpublished data). Taken together, these studies show the importance of the protein lesion size and conjugation site on DNA in defining their biological consequences.

1.5. DPC-inducing agents studied in this thesis

1.5.1. Reactive oxygen species

Reactive oxygen species (ROS) are radical and non-radical oxygen species formed by a partial reduction of oxygen. Examples of ROS include superoxide anions (O_2^-), hydrogen peroxide (H_2O_2), and hydroxyl radicals ($^{\bullet}OH$). ROS can be produced endogenously in cells during aerobic cellular respiration,[338-340] immune response,[341, 342] and inflammation,[35] and after exposure to exogenous agents such as cigarette smoke and urban air,[343, 344] alcohol consumption,[345, 346] and Cr(VI) exposure.[347] ROS production is necessary for cellular survival, as it plays a crucial role in protection against pathogens[348] and cellular signaling pathways such as the phosphoinositide 3-kinase (PI3K) pathway,[349-351] mitogen-activated protein kinase (MAPK) cascades,[352-354] and the NF-κB pathway.[355] However, ROS can also damage cellular biomolecules including DNA and proteins, leading to the direct covalent attachment of proteins to DNA to form DPCs.

1.5.2. ROS-induced DPCs

Some ROS-induced DNA lesions can undergo secondary reactions with nucleophilic sites of proteins, leading to the formation of DPCs.[263, 356] For example, 5-amino-3-β-(2-deoxy-D-ribofuranosyl)-3*H*-imidazo[4,5-d][1,3]oxazin-7-one (2'-deoxyoxanosine, dOxo) lesions are formed from 2'-deoxyguanosine in DNA upon exposure to reactive nitrogen species. The *O*-acylisourea functionality of dOxo is susceptible to nucleophilic attack by protein side chains, forming an amide bond between the protein and the DNA.[357] Interestingly, histone and HMG proteins, which are highly abundant in eukaryotic chromosomes, have relatively low reactivity towards oxanosine as

compared to DNA glycosylases involved in base excision repair, probably a result of damage-specific protein-DNA interactions.[357, 358] Similarly, oxidation of 2'-deoxyribose residues of DNA can lead to the formation of inherently reactive abasic sites. For example, oxidation at the C'-1 position of the 2'-deoxyribose sugar can lead to the formation of 2-deoxyribonolactone, which can form amide conjugates to proteins such as *Escherichia coli* endonuclease III (Endo III) and DNA polymerase β (Pol β).[359,360]

DNA-protein cross-links can also form directly via one electron oxidation of DNA bases. One electron oxidation of guanine can form a radical cation that can react with lysine side chain to yield N^{ε}-(guanin-8-yl)-lysine cross-links.[361] Similarly, hydrogen abstraction from the 5'-methyl group of thymidine yields 5-(uracilyl)methyl radical, which can undergo a radical addition reaction to the C-3 position of a neighboring tyrosine molecule to form thymidine-tyrosine cross-links.[362, 363]

1.5.3. ROS-induced DPC formation following myocardial infarction/reperfusion

One biological event that stimulates a massive and sudden influx of ROS in the heart is myocardial infarction/reperfusion. Myocardial infarction (MI, heart attack) occurs when the blood flow is decreased or stopped in a cardiac artery, restricting the delivery of oxygenated blood to the affected myocardium. In the United States alone, nearly 790,000 individuals suffered a MI in 2016.[364, 365] Current MI therapies are based reperfusion, or the re-opening of the blocked artery, returning blood flow to the damaged myocardium. However, the sudden influx of oxygenated blood is associated with a rapid influx of ROS to the already damaged myocardium.[366-368]

During the myocardial infarction, the lack of oxygen switches the cells to anaerobic respiration mode.[367, 369] Reperfusion-induced oxygen return triggers reactivation of the electron transport chain, producing ROS that can permeate from the mitochondria to other cellular components including the nucleus.[367, 369] Furthermore, neutrophils accumulating in the infarcted myocardium mediate tissue damage release matrix-releasing degrading enzymes and ROS for downstream signaling.[370-372] Neutrophils have been determined to be the primary source of oxygen radicals generated during MI/reperfusion, reaching peak concentrations only 10 min after the reperfusion.[371] Taken together, rapid influx of ROS following MI is expected to oxidatively damage DNA, potentially leading to the formation of covalent DPCs. However, ro our knowledge there are no previous reports of DPC formation following ischemia-reperfusion injury.

1.5.4. Phosphoramide mustard and nornitrogen mustard

Cyclophosphamide (CP) is a nitrogen mustard drug used as first-line chemotherapy treatment for lymphoma,[373, 374] leukemia,[375] myeloma,[376, 377] breast and brain cancers,[378] and in bone marrow transplantation.[379, 380] CP is administered as a prodrug and must be first activated by the liver cytochrome P450 enzymes CYP2B5 and CYP3A4 to yield 4-hydroxy cyclophosphamide (4-OH-CPA).[381, 382] The metabolite 4-OH-CPA spontaneously equilibrates with the ring-open aldophosphamide form, which can enter and accumulate within cells. Aldophosphamide then spontaneously undergoes β-elimination to yield the therapeutically active metabolite, phosphoramide mustard (PM), and the byproduct acrolein. Under physiological conditions, PM spontaneously dephosphoramidates to form another DNA alkylating agent, nornitrogen mustard (NOR).[383-385]

Both PM and NOR can modify the N-7 position of guanine of DNA to yield N-(2-chloroethyl)-N-[2-(7-guaninyl)ethyl] amine, N-(2-hydroxyethyl)-N-[2-(7-guaninyl)ethyl] amine, and *N,N-bis*-[2-(7-guaninyl)ethyl] amine adducts,[383-385] and also produce covalent DNA-protein conjugates. Prior to my investigation, the identities of the proteins participating in PM-induced DPC formation were not known, and the absolute numbers of DPCs induced by nitrogen mustard treatment were not determined.

1.6. Characterization of DPCs

In order to characterize DPC formation, identify protein constituents of DPC, and quantify DPC formation, several biophysical and mass-spectrometry-based techniques have been developed. Early biophysical techniques such as gel shift assays, comet assays, and alkaline elution methods allowed for the confirmation of covalent DPCs and relative quantification. Recently, chemical separation techniques have been developed to isolate covalent-DPCs from unbound chromosomal DNA and proteins respectively. Furthermore, advances in mass spectrometry has allowed for the identification of the protein constituent, the attachment site of the proteins, and the absolute quantification of DPC formation.

1.6.1. Biophysical identification techniques

The simplest technique to confirm DPC formation *in vitro* is the electrophoretic mobility shift assay (EMSA).[334, 386, 387] In this approach, radiolabeled DNA duplexes are incubated with a protein of interest in the presence of a cross-linking agent. The reaction mixture is resolved by denaturing gel electrophoresis and visualized by autoradiography. Under these conditions, any non-covalently bound proteins will dissociate from DNA, while covalent DPCs will remain intact and can be visualized as reduced mobility bands on the gel. Our laboratory has successfully demonstrated DPC formation between specific proteins and nitrogen mustards,[331] 1,2,3,4-diepoxybutane (DEB),[388] and cisplatin.[289] Although EMSAs can easily confirm DPC formation *in-vitro*, they cannot replicate the physiological conditions of the nucleus or provide structural information.

The first reports of DNA-protein cross-link identification in cells utilized biophysical approaches such as alkaline elution methodology.[389] In this method, DNA was

radiolabeled with ^{3}H-labeled thymidine in cell culture. Treated and control cells were lysed onto a polyvinyl filter and washed with the lysis solution to remove intact proteins and lipids.[389] Free DNA was then eluted from the filter by washing with an alkaline solution, leaving the covalent DPCs trapped on the filter.[329, 389, 390] Treatment with proteinase K digested the protein constituent of DPCs, allowing the cross-linked DNA to be eluted through the filter.[329, 390] Relative quantitation of DPCs can be achieved by comparing on the amounts of radioactive DNA eluting from the filter before and after proteinase K digestion.

DPC formation in cells can also be detected by single-cell gel electrophoresis (SCGE), better known as the comet assay.[304, 391-395] Following exposure to cross-linking agent, cells are irradiated to induce strand breaks. Cells are suspended on agarose covered slides, lysed under alkaline conditions, and subjected to electrophoresis. DNA is visualized on the slide by staining with a fluorescent dye. The DNA fragments induced by γ-irradiation will travel through the agarose gel faster than unmodified chromosomal DNA, forming a "tail", while the presence of DPCs will decrease the length of the comet's tail. DPC formation can be confirmed by pretreating the samples with proteinase K and observing a recovery of the comet tail. The length of the comet's tail can therefore be used to estimate the amount of DPC formation. Using the comet assay, Oliver Merk and Gunter Speit were able to show a dose-dependent increase in FA-induced DPC formation from 0 to 500 μM FA in V79 chinese hamster cells.[304] Although both the EMSA and comet assay can identify and relatively quantify DPC formation in cells, these techniques cannot separate DPCs from unmodified chromosomal DNA without destroying the protein

constituent, eliminating any opportunity to identify the protein constituents or the site of cross-linking.

1.6.2. Isolation of DPCs from cells

The first developed method to isolate DPCs formed in cells from unmodified chromosomal DNA was the SDS/K^+ precipitation method developed by Zhitkovich and Costa (**Scheme 1-7**).[396] In this approach, cells are lysed in the presence of SDS and heated to 65 ^oC, allowing for protein denaturation and dissociation from DNA. SDS uniformly binds to the proteins and peptides, yielding a net negative charge on the complex.[397] The lysate is then treated with potassium chloride, and the solution is passed through a pipette tip to uniformly shear DNA. The potassium cations bind to the SDS molecules, creating a net-neutral complex that precipitates DPCs out of solution upon cooling. Zhitkovich and Costa used this assay to detect DPCs induced by chromium, cisplatin, and formaldehyde in CHO cells.[396] This method efficiently separates DPCs from free DNA, allowing for more accurate quantitation. However, it cannot distinguish between proteins binding to DNA non-covalently and those forming covalent DPCs, as both will form SDS/K^+ complexes during the workup.

To ensure that only proteins originating from DPCs are detected, unbound proteins and peptides must be separated from covalent DPCs by removing any non-covalently bound proteins. One of the most common techniques for DNA isolation from cells and tissues is phenol-chloroform or phenol-isoamyl alcohol extraction (**Scheme 1-7**).[398] We have adopted this methodology to isolate DNA containing covalent DPCs from cells treated with DEB,[399] nitrogen mustards,[286-288] and cisplatin.[289] When nuclear lysates from

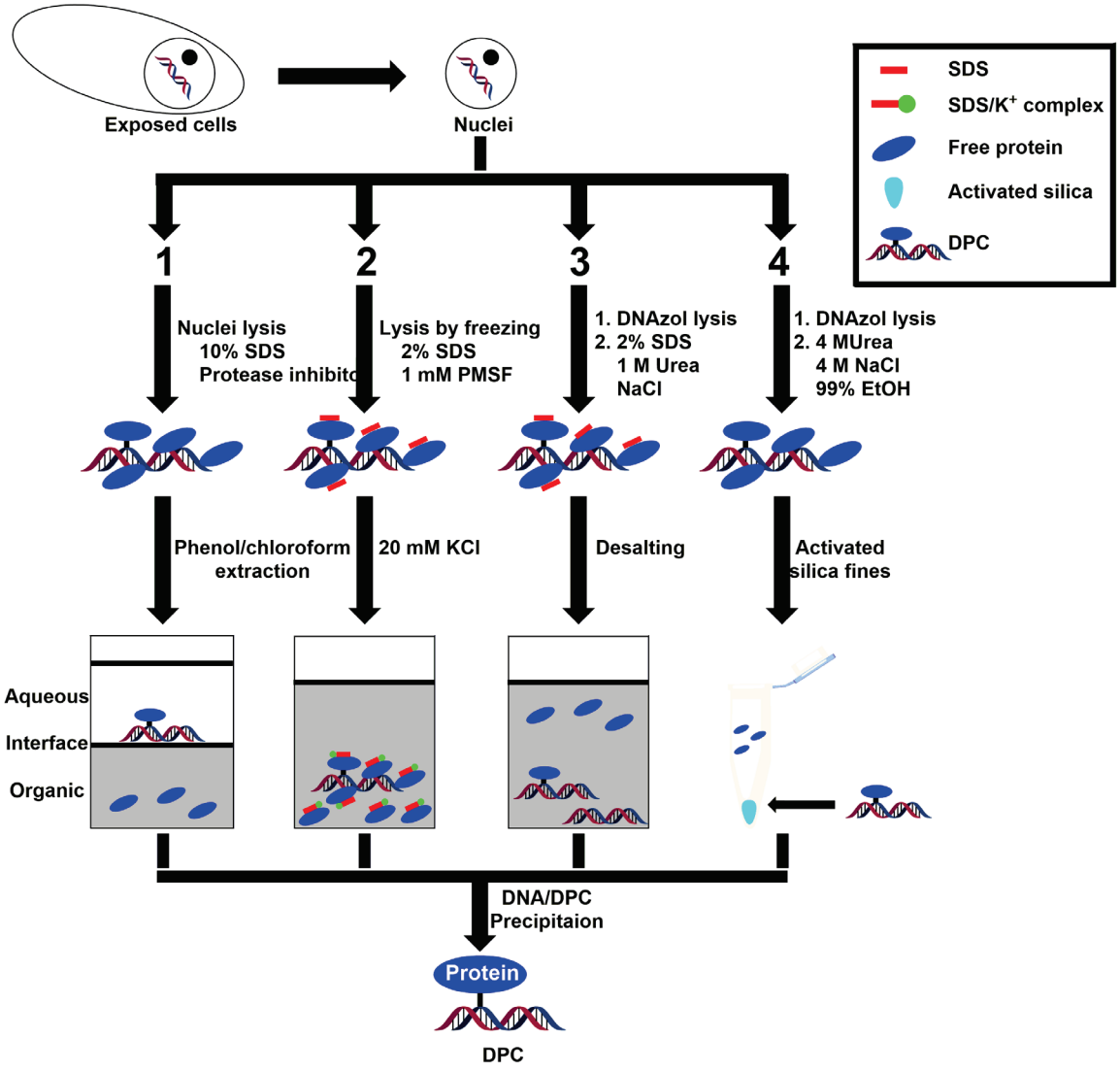
treated cells are mixed with an equal volume of phenol-chloroform, three distinct layers form upon centrifugation. The bottom organic solvent layer contains free proteins, while the top aqueous layer contains chromosomal DNA. The middle layer/interface contains covalent DPCs. This partitioning allows for proteins covalently attached to DNA to be separated from the unbound nuclear proteins. After the top two layers are extracted with organic solvent several times, DNA containing covalent DPCs can be precipitated out of solution with cold ethanol for further analyses via gel electrophoresis or MS-based proteomics. Although DNA and DPCs will be separated from free proteins, there is the potential to contaminate the DNA/DPC aqueous/interface layers with protein-containing phenol during processing, making sample reproducibility more difficult.

Barker *et al* developed the DNAzol-strip and DNAzol-silica extraction methods for DPC isolation from cells (**Scheme 1-7**). Both methods lyse isolated nuclei using the commercially available DNAzol, a guanidine-detergent solution optimized for DNA precipitation. Strong chaotropic agents are used to strip unbound proteins from DNA.[400] DNA containing covalent DPCs is then precipitated with cold ethanol. Although the DNAzol-strip method successfully removes non-covalently bound proteins from DPCs, there is potential for protein degradation due to long processing time in high salt solutions. DNAzol-strip methodology was successfully applied by Barker *et al* to isolate IR-induced DPCs in both CHO and human fibroblast cells[275] and by Yingsheng Wang's group to isolate global FA-induced DPCs from HeLa cells.[401]

In preparation for the DNAzol-silica extraction method, silica fines are activated by heating in nitric acid, and the nuclear lysate is incubated at 37 ºC with 4 M NaCl and 4 M urea to dissociate any unbound proteins. The lysate is then loaded onto the activated

silica fines, where DNA (but not proteins) will be adsorbed. The DNA/DPCs trapped on the silica can be subsequently eluted using 8 M NaOH. As with the DNAzol-strip methodology, there is the potential to degrade proteins on the activated silica during processing. The DNAzol-silica method was successfully used by Prasad *et al* to isolate DPCs formed between PARP-1 protein and abasic sites within DNA.[386]

Scheme 1-7: Schematic representation of methodologies for extracting DNA-protein cross-links from cells *via* phenol/chloroform extraction (**1**), SDS/K^+ precipitation (**2**), DNAzol-strip (**3**), and DNAzol-silica (**4**).



1.6.3. Mass spectrometry based identification of DPCs

Recent developments in the field of mass spectrometry have allowed for the mapping of the cross-linking sites across specific proteins and the specific atomic connectivity of *bis*-electrophiles to DNA and proteins.[402, 403] For example, our laboratory mapped the cross-linking of hAGT and C145A AGT to double stranded DNA using DEB.[388] After hydrolysis of the protein-nucleobase off the DNA backbone and trypsin digestion of hAGT, the tryptic peptides were analyzed by HPLC-ESI$^+$-MS/MS. Two peptides, G^{136}NPVPILIPCHR^{147} and V^{148}VCSSGAVGNYSGGLAVK^{165}, revealed a mass shift of +237 Da that corresponded to the addition of butanediol cross-linked to guanine.[388] CID fragmentation patterns identified Cys^{145} and Cys^{150} as the site of cross-linking.[388] When the same experiment was performed using C145A AGT, the peptide V^{148}VCSSGAVGNYSGGLAVK^{165} displayed the same HPLC retention time and CID fragmentation pattern as the wild type AGT, confirming cross-linking at Cys^{150}.[388]

To confirm DEB-induced cross-linking occurred exclusively at cysteine residues, HPLC-ESI-MS/MS of the total enzymatic digest of the DPC was performed. Following incubation with DEB, the AGT protein was completely digested using by carboxypeptidase Y and proteinase K to yield a single amino acid. HPLC-ESI-MS/MS analysis was set up to monitor the presence of 1-(S-cysteinyl)-4-(guan-7-yl)-2,3-butanediol (Cys-Gua-BD, *m/z* = 359.1), 1-(O-tyrosyl)-4-(guan-7-yl)-2,3-butanediol (Tyr-Gua-BD, *m/z* = 420.1), 1-(N-lysyl)-4-(guan-7-yl)-2,3-butanediol (Lys-Gua-BD, *m/z* = 385.4), 1-(N-arginyl)-4-(guan-7-yl)-2,3-butanediol (Arg-Gua-BD, *m/z* = 413.4), 1-(N-histidyl)-4-(guan-7-yl)-2,3-butanediol (His-Gua-BD, *m/z* = 393.4).[388] Analysis revealed the presence of only Cys-Gua-BD, which co-eluted at the same time as an authentic standard and the same fragmentation

pattern: *m/z* = 359.1 $[M + H]^+$ → *m/z* = 208.1 $[M - Gua + H]^+$, *m/z* = 190.0 $[M - Gua - H_2O + H]^+$. No evidence of Tyr-Gua-BD, Lys-Gua-BD, Arg-Gua-BD, or His-Gua-BD was detected, confirming the atomic connectivity of DEB-induced DPC formation to occur between the N7-position of guanine and the thiol moiety of cysteine. Analogously, our laboratory has performed similar experiments to confirm cross-linking of AGT by nitrogen mustards[331] and cisplatin (Ming *et al*, unpublished data).

1.6.4. Mass spectrometry based proteomics

Recent developments in biological mass spectrometry have allowed great advances in mass spectrometry-based proteomics (MS-proteomics). The new Orbitrap and Q-TOF mass spectrometers provide improved data acquisition speed, high resolution, and excellent mass accuracy (<10 ppm) of peptide fragments, allowing for sensitive and accurate protein identification. These improvements now allow for over 5,000 proteins to be identified in a single proteomics experiment, a number far exceeding the numbers of proteins expected to become cross-linked to DNA. These recent developments have enabled system-wide analyses of complex protein mixtures with excellent sensitivity and specificity, and have recently been applied to identify the protein constituents of DPCs.

The most common approach to identify the protein constituent of DPCs utilizes bottom-up proteomics methodology. In this method, proteins are digested to peptides with specific proteases and identified based on their molecular weights and MS/MS fragmentation patterns, which are compared to theoretical spectra present in spectral databases.[404] To improve digestion efficiency, disulfide linkages of proteins are reduced to thiols using dithiolthreitol (DTT) or *tris*(2-carboxyethyl)phosphine (TCEP), which are

alkylated with iodoacetamide to prevent their oxidation. Proteolytic digestion of proteins generates predictable peptide fragments due to protease-specific peptide bond cleavages. The most common protease used in proteomics is trypsin, which cleaves proteins at arginine and lysine residues.[405] Other commonly used proteases include chymotrypsin (cleaves at carboxyl side of lysine and the aromatic residues tyrosine, tryptophan, and phenylalanine), Lys-C (cleaves at carboxyl side of lysine and arginine or proline), Glu-C (cleaves at the C-terminus of aspartic or glutamic acid and serine), and Arg-C (cleaves at the C-terminus of arginine and proline or lysine).[404]

Typical bottom-up proteomics analyses utilize nano HPLC separation of peptide mixtures coupled to nanospray ionization tandem mass spectrometry analysis, preferably using high resolution mass analyzers such as Orbitrap MS (**Scheme 1-8**).[404] Data is acquired in an automated data-dependent MS/MS mode, where full scan events are followed by MS^2 or MS^3 analyses of the most abundant ions in the spectra. Details regarding mass spectrometers, liquid chromatography separation of peptides, and fragmentation patterns of peptides are beyond the scope of this review and have been reviewed elsewhere.[404, 406]

Mass spectrometry-based proteomics provides an opportunity to conduct global, unbiased analysis of all protein components of DPCs (**Scheme 1-8**). This approach has been successfully used to investigate DPC formation by many cross-linking agents including ionizing radiation,[274, 275, 407] reactive oxygen species, cisplatin,[335] nitrogen mustards,[331] diepoxybutane,[388] and formaldehyde.[401] In order to utilize bottom-up proteomics to identify DPCs, the protein constituent of a DPC must be released from the DNA backbone. Because many *bis*-electrophiles alkylate DNA at the N7 position of

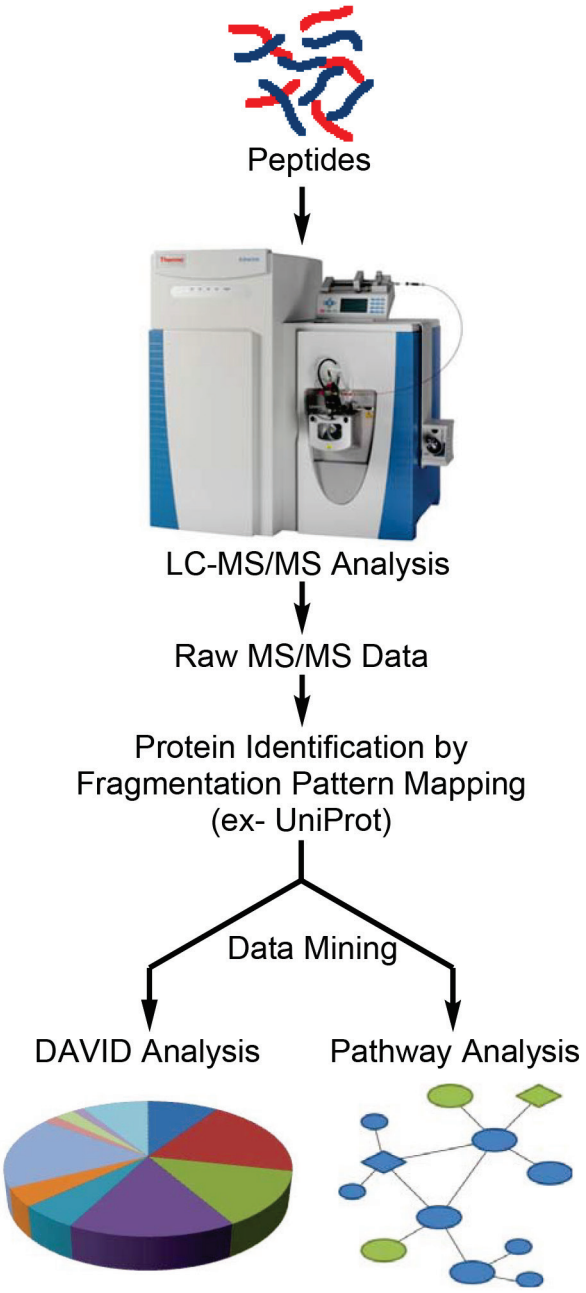
guanine, this creates hydrolytically labile DNA lesions susceptible to depurination and the formation of abasic sites. Upon heating to 90° C, such DPCs can be quantitatively released from the DNA backbone in the form of protein-guanine conjugates. We have recently shown that this methodology can be modified to allow for analyses of hydrolytically stable DPCs such as those induced by reactive oxygen species by incorporating a DNA digestion step (see Chapter 3 of this Thesis). DNA can be digested to single nucleosides in the presence of nuclease enzymes to generate protein-nucleoside conjugates suitable for proteomics analyses.

Following their release from DNA, proteins participating in cross-linking can be resolved by SDS-PAGE and visualized by staining with mass spectrometry-compatible stains such as Coomasie brilliant blue.[408] Individual protein bands or gel regions corresponding to specific molecular weight ranges can then be excised from the gel, allowing for initial fractionization of the proteins mixtures. The excised gel pieces can then be prepared for trypsin digestion followed by extraction of the resulting tryptic peptides from the gel matrix. As an alternative to gel-based analyses, filter-aided sample preparation (FASP) techniques can be used to simplify sample preparation for MS-based proteomics.[409] This protocol performs protein capture and digestion steps on the same size exclusion filters. This method allows for high concentrations of detergents (i.e. 8 M urea) to be used to ensure universal solubilization of proteins. Following reduction and alkylation steps described above, detergents can be washed off the filter by centrifugation, while the proteins remain trapped on the filter. Finally, proteins can be digested to peptides using trypsin, and the resulting peptides are eluted and collected for MS analysis. For complex

samples, peptides generated by FASP can be pre-fractionated using offline HPLC, and each fraction is separately analyzed by HPLC-ESI^{+}-MS/MS.

To allow for protein identification, experimentally observed peptide mass and MS/MS fragmentation data are entered into a software package such as Proteome Discoverer (Thermo-Fisher), where the fragmentation patterns are compared to spectra present in the database (**Scheme 1-8**). Identified proteins are matched to their corresponding gene ID, creating a gene list. From here, various data mining tools can be used to probe the functions of the identified proteins. Singular enrichment analysis tools such as DAVID (Database for Annotation, Visualization, and Integrated Discovery) can be used to extract the molecular and biological functions of the identified protein.[410] While large gene lists can provide an overwhelming amount of information due to the linearity of the output terms, these terms can then be condensed into relevant categories allowing enriched biological and functional themes to be detected. For example, terms such as apoptosis, programmed cell death, and regulation of apoptosis, etc. may be condensed under a single term, apoptosis, in order to simplify the data output.[410] Another method of data mining that can be employed is the use of pathway analysis software, such as Ingenuity Pathway Analysis (IPA). Here, the gene list is inputted into the software where relationships and canonical pathways involved can be elucidated through comparisons with an extensive collection of published experimental data.[411] These are just two examples of the numerous data mining techniques that can be utilized to extract biological meaning from the extensive gene lists produced by proteomics experiments.

Scheme 1-8: Schematic of protein identification from mass spectrometry-based proteomics data analysis



Protein identities obtained from LC-MS/MS can be confirmed using immunological detection with specific antibodies, also referred to as immunoblotting. Incubation with a conjugated secondary antibody allows the visualization of these specific proteins using colorimetric or chemiluminescence detection. In preparation for immunoblotting, DPCs are digested to protein-guanine or protein-nucleoside conjugates, which can then be resolved by one-dimensional or two-dimensional gel electrophoresis. In two-dimensional gel electrophoresis, proteins are resolved by two differing properties. For instance, proteins can be first separated by isoelectric point using a pH gradient gel and then transferred to a PAGE gel to further separate the proteins based on mass. Following either gel electrophoresis technique, the gels can be incubated in the presence of specific antibodies for identification of individual proteins. Although immunological techniques are selective and sensitive, they can only detect several proteins at a time, requiring prior knowledge of protein identities.

1.6.5. Quantitative proteomics of DPCs

Although traditional bottom-up mass spectrometry-based proteomics methodologies can help identify protein constituents of DPCs, they are unable of quantifying DPC amounts following exposure to cross-linking agents. Background levels of DPCs are present in all living cells due to endogenous exposure to reactive oxygen species, products of lipid peroxidation, and due to inadvertent trapping of DNA topoisomerases, polymerases, and repair proteins on their DNA substrates.[29, 30, 412, 413] In order to distinguish between background and exposure-mediated cross-linking, the amounts of DPCs in control and treated cells must be compared. Recently developed

quantitative proteomics techniques use stable isotope labeling of proteins or peptides to allow for relative quantitation of proteins across multiple samples/experiments.[414, 415] These quantitative proteomics protocols have been recently applied to studies of DNA-protein cross-linking, allowing for better understanding of exogenous versus induced DPC formation, as well as their persistence and repair over time.

One of the most powerful techniques of quantitative proteomics is stable isotope labeling with amino acids in cell culture (SILAC). In this approach, isotopically labeled arginine and lysine are added to the cell culture media, leading to their incorporation into all cellular proteins during translation.[414, 416] Adding two labeled amino acids ensures that every peptide from a proteolytically digested protein will contain at least one isotopic label.[416] In a typical SILAC experiment, experimental manipulation such as exposure to a cross-linking agent is performed with the isotopically labeled "heavy" cells, while the control condition is performed with the normal "light" cells. Following treatment, proteins from light and heavy cells are combined and subjected to traditional mass spectrometry-based proteomics workflow. Since the physiochemical properties, such as ionization potential, fragmentation, and ESI MS signal response are the same for isotopically labeled and unlabeled peptides, their relative HPLC-MS/MS signal intensities allow for relative quantitation across samples, enabling the detection of quantitative changes in protein levels. To confirm the quantitative results, reverse labeling can be employed with heavy cells serving as controls this time.

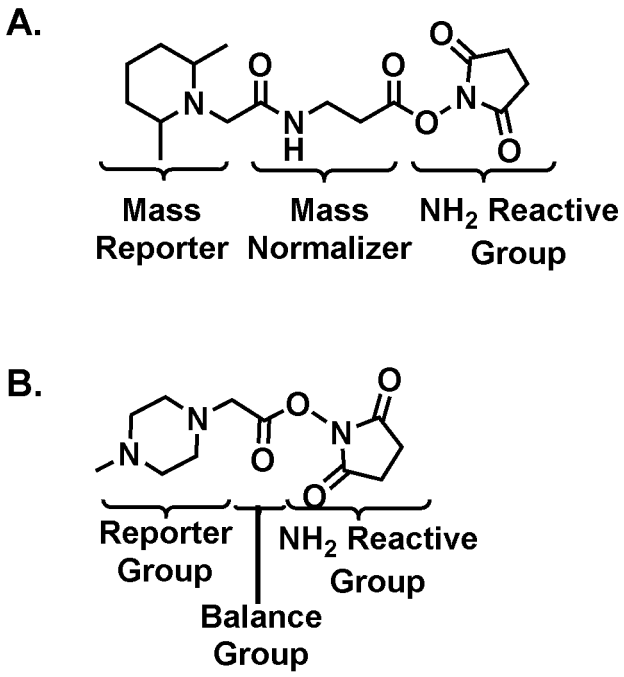
To our knowledge, SILAC has not yet been used to quantify DPC formation following exposure to environmental and therapeutic *bis*-electrophiles. However, DNA-protein cross-linking by formaldehyde in combination with SILAC has been used to

measure changes in DNA-protein binding. For example, Byrum *et al* combined chromatin affinity purification with SILAC mass spectrometry to quantify changes in protein binding and histone post-translational modifications to the *GAL1* loci under "transcriptionally silent" or "active" states in *S. Cerevisiae.*[417] These authors introduced a LexA binding site upstream of the *GAL1* promoter in a strain of *S. Cerevisiae* constitutively expressing a LexA fusion protein.[417] The mutant strain of yeast was grown in regular light media, while the wild type strain was grown in heavy media. Both strains were then exposed to glucose (to repress transcription) or galactose (to activate transcription), followed by FA treatment to cross-link bound proteins to DNA. DNA shearing and ChIP purification using a LexA antibody allowed for purification of the *GAL1* loci along with any trapped proteins. SILAC-based proteomics has enabled these authors to identify and quantify the proteins interacting with "silent" or "active" chromatin, as well as to eliminate any non-specific binders which manifested themselves by having 1:1 ratio of light and heavy peptide signals. A similar approach was adopted by Soldi and Bonaldi [310] to observe global changes in protein binding at transcriptionally active or silent regions of chromatin. Following treatment with FA to induce cross-linking of DNA-binding proteins to DNA, regions of chromatin were immunoprecipitated using antibodies against histone H3 with a trimethylated lysine 9 (transcriptionally repressed)[418] and histone H3 with a trimethylated lysine 4 (transcriptionally active) [419], allowing for the relative quantification of proteins binding to various regions of chromatin. These examples suggest that SILAC could also be used for quantitation of DPCs induced by endogenous and exogenous agents. However, SILAC requires *in vitro* labeling of proteins with unnatural isotopes and thus is limited to

cell culture experiments. Furthermore, multiplexing experiments are not practical due to limited availability of isotopically labeled cell culture media and cost considerations.

Another powerful methodology of quantitative proteomics involves chemical labeling of reactive residues of tryptic peptides. Two of the most popular chemical labeling agents are tandem mass tags (TMT, available from Thermo Fisher Scientific) (**Figure 1-12A**) and isobaric tags for absolute and relative quantification (iTRAQ, sold by Sciex) (**Figure 1-12B**). Both TMT and iTRAQ tags contain a mass reporter region, a mass normalizer section, an amine-reactive N-hydroxysuccinimide group, and the same number of isotopic labels differentially distributed throughout the molecule. These tags react with and label lysine and arginine residues and the N-terminal amines of tryptic peptides. Experimental samples and controls are labeled with isotope unique tag, and the labeled peptides are mixed together prior to analysis. Upon MS/MS or MS^3 fragmentation of peptides during mass spectrometry analyzes, characteristic product ions are observed in the spectra. Since each mass reporter region will have a different number of isotopic labels, peptides from differentially labeled samples can be distinguished from each other, thus providing a characteristic signal for peptides originating from each sample/experimental condition. Furthermore, reporter ion intensities can be used for relative quantitation of each protein across samples. Commercially available kits contain four, six, or eight unique labels, allowing for ready multiplexing. I utilized TMT labeling to quantify DPC formation in cardiomyocytes following myocardial infarction/reperfusion as compared to healthy myocardium using 6-plex TMT methodology, which is described in detail in chapter 3.

Figure 1-12: Chemical structures of commercially available chemical labels Tandem Mass Tags (TMT) (**A**) and isobaric tags for relative and absolute quantitation (iTRAQ) (**B**) used in quantitative proteomics



1.7. Mechanisms of DPC repair

The mechanisms by which covalent protein-DNA complexes are recognized as DNA damage and removed from genomic DNA in cells and tissues are incompletely understood. DPCs that are produced by the enzymatic entrapment of topoisomerases 1 and 2 (TOP1 and TOP2) can be reversed by specific tyrosil-DNA phosphodiesterases TDP1 and TDP2, respectively.[420] These enzymes hydrolyze the covalent bond between the topoisomerase's active site residue and the DNA, releasing the protein.[420] For all other DPC structures, such as those formed endogenously by ROS exposure or sequential cross-linking by *bis*-electrophiles, the repair mechanism remains controversial.

In the yeast, Stingele *et al* identified the metalloprotease Wss1 as a DPC-specific protease that could relieve replicative stalling by acting directly on the covalently attached protein.[421]Yeast mutants lacking Wss1 exhibited an accumulation of FA-induced DPCs and topoisomerase-mediated DPCs, lead to yeast accumulating in the late-S/G2 phase of cell cycle.[421] Furthermore, yeast lacking Wss1 exhibited an increase in FA-induced recombination and gross chromosomal recombination.[421] This is consistent with a model where DPCs cause replication fork stalling, leading to cleavage by endonucleases and one-ended DNA double-strand breaks, which in turn lead to recombinational repair.[422]

When the roles of canonical DDR pathways in DPC repair were investigated in mammalian cells, varied conclusions were achieved. Zhitkovich et al. examined toxicity and repair of DPCs induced by Cr(VI) and formaldehyde in human cells.[324, 423] These researchers found that pharmacological inhibition of nucleotide excision repair increased DPC formation, while inhibition of the proteasome activity abolished DPC repair, suggesting that proteosomal degradation of DPCs initiates their repair by NER.[309] Nakano

et al reported that proteins smaller than 14 kDa could be removed by NER *in itro,* but larger proteins were not NER substrates.[424, 425] Additionally, DNA-peptide cross-links produced from protease-digested DPCs have been identified as substrates for nucleotide excision repair *in vitro*.[309, 426, 427] Reardon et al. and Baker et al. demonstrated that DPCs containing T4-pdg (16 kDa) and *Hhai I* DNA methyltransferase proteins (HDnmt, 37 kDa) could not be excised by human nucleotide excision repair proteins *in vitro*, but smaller polypeptide-DNA cross-links were efficiently removed by NER.[426, 427]

Alternatively, Nakano *et al* has proposed homologous recombination as the major DDR pathway that contributes to DPC removal. They were unable to detect polyubiquitination of extracted FA-induced DPCs by western blot, which would be required for proteasome degradation.[424] Furthermore, Chinese hamster cells deficient in RAD51 (HR-deficient) were two to three magnitudes more sensitive to formaldehyde and 5-aza-2'-deoxycytidine (azadC) compared to NEHJ or NER-deficient cells, and recovery of RAD51 eliminated this observed sensitivity to the drugs.[424] Finally, HR-deficient cells treated with FA or azadC were found to accumulate double strand breaks via alkaline comet assay (see section 1.5.1).[424] Taken together, data of Nakano *et al* suggest that DPC formation leads to replication fork breakage at the site of DPC formation and thus initiates HR to alleviate the stalled replication fork.[424]

Recent studies have focused on the role of the metalloprotease DVC1/Spartan (SPRTN) in digesting the protein constituent of a DPC to allow for their repair. SPRTN is considered as a mammalian homolog to the yeast Wss1, as SPRTN shares both similar domain organization and a common evolutionary origin.[428] SPRTN knockouts are embryonic lethal, while hemizygous SPRTN rats experience rapid ageing, genome

instability, kyphosis, and early development of cataracts.[429] Enzymatic digestion of the protein constituent of a DPC was first confirmed by Duxin *et al.*[430] They incubated DPC-containing plasmid and the corresponding plasmid control with *Xenopus* egg extract, followed by restriction enzyme digestion to yield a 165-nucleotide fragment containing the DPC.[430] Gel electrophoresis of the plasmid digestions yielded one band for the control and multiple bands for the DPC-containing plasmid, indicating the protein constituent was enzymatically digested.[430] Considering the evolutionary similarities with Wss1, the authors proposed that Spartan acts as the DPC-specific protease.[430]

Using *Xenopus* egg extract and DPC-containing plasmids, Duxin *et al* also examined DPC repair mechanism that couples DNA replication and protein digestion to adduct bypass. Two-dimensional gel electrophoresis of control and DPC-containing plasmids incubated in the extract revealed that a DPC-containing plasmid could not be replicated.[430] When the control plasmid was incubated in the extract for seven minutes, only one linear molecule was observed, indicating the plasmid was completely replicated. However, DPC-containing plasmids migrated much slower and in a wider band, which was consistent with fork convergence and stalling at the lesion.[430] Furthermore, when the DPC-containing plasmid was incubated in extract with nonfunctional DNA polymerase ζ (mutations to subunits Rev1 and Rev7), extension product of the adduct-containing strand was greatly diminished.[430] In total, the authors concluded that the collision of the replisome/CMG helicase with a DPC located on the DNA leading strand triggers proteolytic degradation of the protein constituent of DPC, and the resulting peptide-DNA adduct is ultimately bypassed DNA polymerase ζ.[430]

The molecular mechanism and regulation of Spartan-induced protein digestion has been further elucidated in the past year. Using isolated Spartan, Stingele *et al* observed proteolytic digestion of the DNA-binding proteins histone H1, H2A, H2B, H3, and Hmg1 in the presence of ssDNA.[431] When Spartan was incubated with the same proteins in the presence of dsDNA, only autocleavage of Spartan was detected.[431] Furthermore, no Spartan activity was detected when GFP and BSA, proteins without DNA binding regions, were incubated with ssDNA.[431] Taken together, Spartan's protease activity was determined to be controlled by a DNA switch, where dsDNA allows for autocleavage, while ssDNA also allows for proteolytic digestion of DNA-bound proteins.

In cells, Spartan exists in two forms: unmodified and monoubiquitinated, with a portion of Spartan always bound to chromatin.[432] Stingele *et al* observed only unmodified Spartan bound to chromatin by western blot, while both unmodified and mono-ubiquitinated Spartan were observed in the soluble protein extract.[431] When cells were treated with FA, a noticeable shift from monoubiquitinated Spartan to chromatin-bound unmodified Spartan was observed over time, suggesting that DPC induction by FA triggered ubiquitin removal from Spartan.[431] Stingele *et al* concluded that Spartan localization is controlled by ubiquitination, with deubiquitinating leading to Spartan binding to chromatin.[431] Specifically, Spartan will bind to ssDNA regions created by a DPC, bringing Spartan into proximity of its substrate and allowing for proteolytic digestion of the protein constituent, despite Spartan's low affinity for proteins in general.[431]

Additional research has coupled Spartan's protease activity with the inhibition of DNA replication in mammalian cells. Morocz *et al* utilized a DNA fibre assay of FA-treated cells to show Spartan is required for an immediate bypass of DPCs.[433] The speed

of replication with or without DPC induction was determined by first pulse-treating cells with iododeoxyuridine (IdU, red label), followed by simultaneous incubation with FA and bromodeoxyuridine (BrdU, green label) and immunoblotting against the nucleotide analogues to visualize newly replicated DNA tracks. In FA-treated Spartan-depleted cells, the length of newly formed DNA tracks was significantly decreased compared to FA-treated wild-type cells, indicating a strong inhibition of bypass across the FA-induced damage in the absence of Spartan.[433] When the DNA fibre assay was performed with cells expressing Spartan with a mutated protease domain (SprT), DNA replication was inhibited to the same degree as Spartan depletion.[433] These results are consistent with the proposed DPC repair model where Spartan degrades the protein constituent to a peptide fragment that can be bypassed by translesion synthesis.[433]

Although significant progress has been made toward elucidating the mechanism of DPC repair, many questions still remain. For example, the shared contributions of NER and HR towards removing DPCs of varying size have yet to be fully elucidated. Furthermore, no direct evidence of Spartan proteolytically digesting a model DPC *in-vitro* or digesting chromatin-bound DPCs *in vivo* has been reported. Similarly, the contributing effects of cell cycle stage, identity and/or size of the protein, and molecular shape of the linker unit needs to be further investigated.

1.8. Thesis statement

DNA-protein cross-links are ubiquitous, super-bulky adducts formed after the exposure to many physical and chemical agents. If left unrepaired, DPCs will interfere with crucial cellular processes such as DNA replication and transcription. In order to better understand the biological consequences of DPCs induced by *bis*-electrophiles or ROS, the identity of the protein constituents, the atomic connectivity, and the extent of DPC formation need to be elucidated.

Towards these goals, mass spectrometry-based proteomics was utilized to characterize PM- and NOR-mediated DNA-protein cross-linking human fibrosarcoma (HT1080) cells at cytotoxic concentrations (Chapter 2). PM is the active metabolite of the antitumor drug cyclophosphamide, and PM-induced DPCs are expected to contribute to the biological effects of cyclophosphamide treatment. Over 130 proteins were found to be covalent trapped to DNA only after PM-exposure, including those involved in transcriptional regulation, RNA splicing/processing, chromatin organization, and protein transport. To structurally characterize PM- and NOR-induced DPCs, a HPLC-ESI$^+$-MS/MS assay was developed to analyze the proteolytic digests of DPC-containing DNA from NOR-treated. A concentration-dependent formation of *N*-[2-[cysteinyl]ethyl-*N*-[2-(guan-7-yl)ethyl]amine (Cys-NOR-N7G conjugate was observed, confirming that it cross-links cysteine thiols of proteins to the N7 position of guanines in DNA

Furthermore, we began investigating the DDR pathways that contribute to DPC repair. Cell cytotoxicity experiments revealed that NER-deficient (XPA-and XPD-deficient) and HR-deficient (FANCD2-deficient) fibroblast cell lines were twice as sensitive to PM-treatment compared to isogenic controls. To investigate DPC repair over

time, the cell lines were treated with an equal concentration of PM for 3 hours, followed by incubating an additional 0, 4, 8, or 24 hours to allow the cells to repair PM-induced damage. Analysis by the developed LC-MS/MS assay revealed XPA-deficient and XPD-deficient cells had the highest Cys-NOR-N7G levels at every time point, indicating these cells had the least amount of DPC repair over time. Conversely, there was no difference in Cys-NOR-N7G levels between HR-deficient cells and their isogenic controls. Taken together, we concluded that NER may be involved in the repair of PM- and NOR-induced DPCs, and NER contribution to DPC repair needs to be further investigated.

The second goal of my DPC project was to characterize and quantify endogenous DPC formation after exposure to ROS. Specifically, I utilized an animal model of myocardial infarction/reperfusion (LAD ligation/reperfusion) to determine whether the known influx of ROS following reperfusion (ie. reperfusion injury) would lead to an increase in ROS-induced DPC formation. One mechanism of ROS-induced DPC formation involves a hydroxide anion abstracting a hydrogen from the methyl group of thymidine, which allows a free radical addition to a tyrosine side chains of proteins. I synthesized thymine-tyrosine cross-links (dT-Tyr) and developed a sensitive isotope dilution tandem mass spectrometry assay for their detection in cells and tissues. Quantitative analysis of DPCs extracted from rat cardiomyocytes after LAD ligation/reperfusion revealed an increase in DPC formation.

Quantitative proteomics experiments (TMT labeling) were utilized to both identify ROS-induced DPCs and relatively quantify DPC formation after MI/reperfusion compared to a sham surgery. A total of 90 proteins participating in ROS-induced DPC formation were identified, including ROS scavengers, contractile proteins, and regulators of apoptosis.

TMT quantitation revealed that 80 of the DPC proteins were at least 1.2-fold more abundant after LAD ligation/reperfusion compared to sham surgeries. Global proteome changes were far less pronounced, but did exhibit increased expression of mitochondrial proteins required for aerobic respiration and biomarkers of sarcomere breakdown following ischemia/reperfusion injury. Overall, these quantitative results are consistent with a model where reperfusion injury and the sudden influx of ROS leads to an increase in DPC formation, and these lesions may contribute to long-term cardiac injury.

Another goal of my DPC projects was to investigate the role of the metalloprotease Spartan in ROS-induced DPC repair (Chapter 4). Spartan is believed to proteolytically digest the protein constituent to a smaller peptide fragment, which could then be excised from DNA by NER. Cell cytotoxicity experiments revealed that floxed hemizygous SPRTN MEF cells were more sensitive to DPC-inducing agents such as hydrogen peroxide, cisplatin, phosphoramide mustard, and 1,2,3,4-diepoxybutane compared to isogenic controls. Furthermore, endogenous DPC formation across different organs in wild-type mice or hemizygous SPRTN-mice was quantified using the developed dT-Tyr nanoLC-MS/MS assay. A small but significant increase in DPC formation was observed in the heart, brain, liver, and kidneys of SPRTN-deficient mice as compared to the wild-type organs. No detectable levels of dT-Tyr were observed in the thymus or spleen of both classes of mice. The increase in dT-Tyr in several tissues from SPRTN-deficient mice compared to wild-type marks the first evidence that the Spartan protease is involved in DPC repair. Furthermore, the lack of detectable dT-Tyr is tissues such as the thymus and spleen indicates that additional proteases must be involved in the digestion and repair of DPCs.

As described in section 1.2, many BD-induced DNA nucleoside adducts have been characterized and observed *in vitro* or *in vivo*. However, ring open N^5-EB-FAPy-dG and N^5-DEB-FAPy-dG had not been previously investigated. The final goal of my thesis was to structurally characterize and quantify EB-FAPy-dG *in-vitro* and in cells (Chapter 5). I first synthesized authentic standards of N^5-EB-FAPy-dG and structurally characterized the adducts by ^{1}H-NMR, HSQC, and HMBC. These authentic standards were used develop an isotope dilution tandem mass spectrometry assay in order to quantify N^5-EB-FAPy-dG formed *in vitro*.

When CTDNA was incubated with an increasing amount of EB (0 – 10 mM) and treated with base, a dose dependent increase in EB-FAPy-dG was observed. To determine the optimal pH for imidazole-ring opening, CTDNA was incubated with 5 mM EB, followed by treatment with tris-buffer at pH 7.5 – 12.0. EB-FAPy-dG was detected in every sample, indicating that EB-FAPy-dG formation could occur at physiological pH. However, EB-FAPy-dG levels were two-fold higher at pH 12, showing that a significant amount of alkylated dG was not subject to ring opening. To investigate EB-FAPy-dG vs EB-Gua II formation *in-vivo*, MEF cells proficient and deficient in the BER enzyme NEIL1 were treated with 500 μM EB for 24 hours. Both EB-FAPy-dG and EB-Gua II levels were increased ~ 2-fold in NEIL1 deficient cells, suggesting that NEIL1 may participate in the repair of EB-FAPy-dG. Taken together, we provided the first evidence of FAPy-dG formation in cells after exposure to the metabolites of 1,3-butadiene.

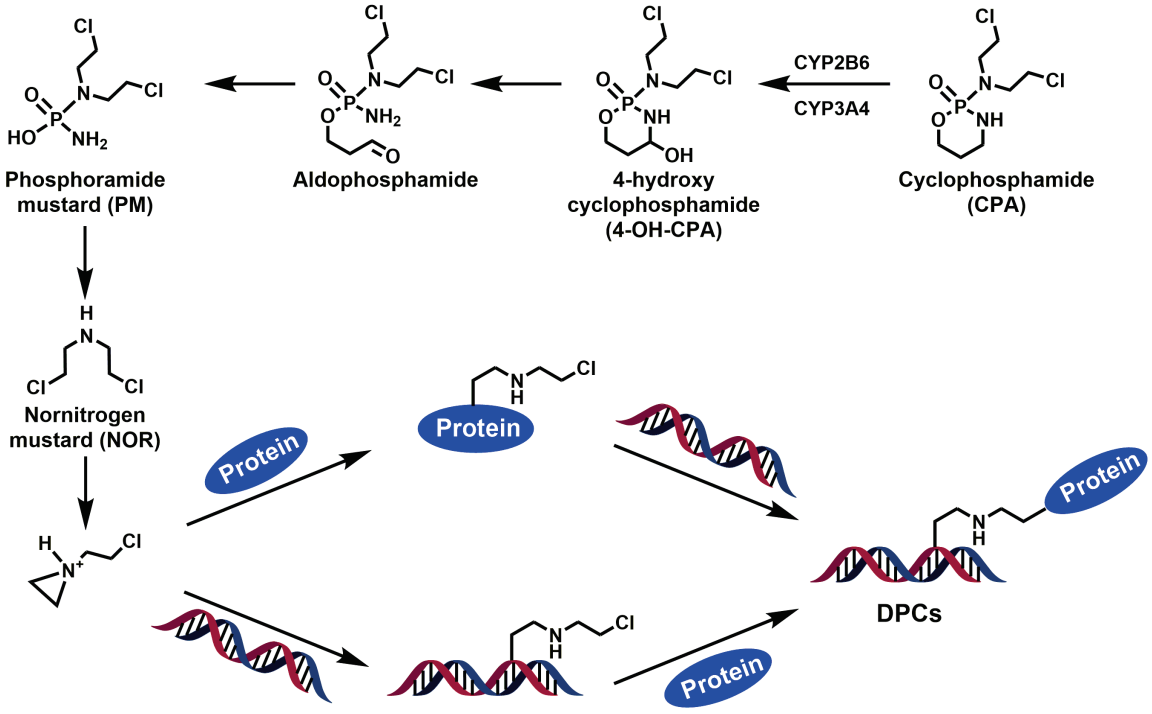
2. Covalent DNA-Protein Cross-Linking by Phosphoramide Mustard and Nornitrogen Mustard in Human Cells

2.1. Introduction

DNA-protein cross-links (DPCs) are bulky DNA lesions that are formed as a result of irreversible trapping of cellular proteins on DNA strands.[262] DPCs can form endogenously or may be induced by exposure to transition metal ions,[281, 434] ionizing radiation,[275] and anti-cancer drugs such as nitrogen mustards,[287, 329, 435] platinum compounds,[334, 436-438] and haloethylnitrosureas.[390] As a result of their size, DPC lesions are thought to interfere with chromatin folding, DNA replication, transcription, and repair, potentially contributing to toxicity and mutagenicity of cross-linking agents.[263]

Phosphoramide mustard (PM, **Scheme 2-1**) is a biologically active metabolite of cyclophosphamide (CP), a chemotherapeutic agent commonly used to treat lymphomas, breast cancers, certain brain cancers, and autoimmune diseases.[326, 378, 439] Under physiological conditions, PM spontaneously dephosphoramidates to form another DNA alkylating agent, nornitrogen mustard (NOR) (**Section 1.5.4, Scheme 2-1**).[383-385] Both PM and NOR can modify N-7-guanine of DNA to yield N-(2-chloroethyl)-N-[2-(7-guaninyl)ethyl] amine, N-(2-hydroxyethyl)-N-[2-(7-guaninyl)ethyl] amine, and *N,N-bis*-[2-(7-guaninyl)ethyl] amine adducts,[383-385] and also produce covalent DNA-protein conjugates (**Scheme 2-1**).

Scheme 2-1: Metabolism of cyclophosphamide to phosphoramide mustard and nornitrogen mustard and the formation of DNA-protein crosslinks.



DNA-protein cross-linking by nitrogen mustards was first observed by Ewig *et al.*[329] These authors utilized the alkaline elution approach to detect DPCs in mouse leukemia cells (L1210) treated with nitrogen mustard.[329] Hansson *et al* used a similar methodology to demonstrate DPC formation in human melanoma cells upon exposure to NOR and melphalan.[330] When compared to other bifunctional lesions such as DNA interstrand and intrastrand cross-links, DPCs accounted for 60 – 70% of total cross-linked lesions.[330] Our group initially investigated DNA-protein cross-linking by nitrogen mustards using a model protein, O^6-alkylguanine DNA alkyltransferase (AGT).[331] Covalent AGT-DNA conjugates formed in a concentration-dependent manner following treatment with mechlorethamine and chlorambucil[331] and involved the N7 position of guanine in DNA and AGT active site cysteine residues, C145 and C150.[331] More recently, a mass spectrometry-based proteomics approach was employed to identify mechlorethamine-induced DPCs in human fibrosarcoma (HT1080) cells.[287] A total of 38 proteins were identified, including those participating in DNA damage response/repair, RNA processing/mRNA splicing, and transcriptional regulation/translation.[287] Unfortunately, many of the cross-linked proteins remained undetected due to the limitations of mass spectrometry methodologies used at the time (ion trap MS).[287]

In the present study, highly sensitive and accurate Orbitrap methodology was used to characterize DNA-protein cross-linking in HT1080 cells treated with PM and NOR. As mentioned above (**Scheme 2-1**), NOR forms upon spontaneous dephosphoramidation of PM, and both mustards produce structurally identical DNA and protein adducts. Proteins covalently bound to chromosomal DNA following PM treatment were isolated using a modified phenol/chloroform extraction methodology developed by our group.[287, 440]

Neutral thermal hydrolysis was used to release PM-induced DPCs from the DNA backbone in the form of guanine-protein cross-links, which were subsequently resolved by SDS-PAGE and identified by mass spectrometry-based proteomics on an Orbitrap Velos mass spectrometer. A total of 134 proteins were found to form DPCs in the presence of PM, including gene products that function in DNA repair, transcriptional regulation, apoptosis, and cell signaling. Isotope dilution HPLC-ESI$^+$-MS/MS analyses of *N*-[2-[cysteinyl]ethyl]-N-[2-(guan-7-yl)ethyl]amine (Cys-NOR-N7G) conjugates with $^{15}N_5$-labeled internal standard were used to quantify DPC formation in NOR-treated cells. Finally, DPC numbers were compared between human wild type fibroblasts and the corresponding cells deficient in nucleotide excision repair (NER) and Fanconi Anemia (FA) repair pathways to gain insight into potential mechanisms of DPC repair in humans.

2.2. Materials and methods

Chemicals and Reagents

Ammonium bicarbonate, ammonium acetate, diepoxybutane, *bis*-(2-chloroethyl)amine hydrochloride (nornitrogen mustard), mechlorethamine, *d,l* -1,2,3,4-diepoxybutane, phenylmethanesulfonyl fluoride (PMSF), Boc-L-cysteine (Boc-Cys-OH), triflouroacetic acid (TFA), leupeptin, pepstatin, aprotinin, methoxyamine, UCN-01, dithiothreitol (DTT), iodoacetamide, chloroform, ribonuclease A, deoxyribonuclease I, and alkaline phosphatase were purchased from Sigma (St. Louis, MO). Phosphodiesterase I and phosphodiesterase II were obtained from Worthington Biochemical Corporation (Lakewood, NJ). UltraPure buffer-saturated phenol was obtained from Invitrogen (Carlsbad,CA). Mass spectrometry grade trypsin was purchased from Promega (Madison,

WI). Proteinase K was obtained from New England Biolabs (Beverly, MA). Cell Lysis Solution and Protein Precipitation solution were purchased from Qiagen (Valencia, Ca). Phosphoramide mustard was obtained from iTT GmbH/Niomech (Bielefeld, Germany).

***N-(2-chloroethyl)-N-[2-(guan-7-yl)ethyl]amine* (*N7G-NOR-Cl*)**

2'-Deoxyguanosine (500 mg, 1.87 mmol) was reacted with nornitrogen mustard (3.34 g, 18.7 mmol) in trifluoroethanol (25 mL) at 37 °C for 72 h under anhydrous conditions. The reaction mixture containing N7G-NOR-Cl was dried under argon, and the resulting solid was washed with 1 mL of anhydrous ether three times to remove unreacted NOR. The presence of N7G-NOR-Cl was confirmed by UV spectrophotometry and mass spectrometry. UV: λ_{max} = 246 nm, λ_{min} = 275 nm (pH 4.9); ESI^+-MS/MS: *m/z* 257.7 [M + H], *m/z* 178.2 $[M + H - NH(CH_2)_2Cl]^+$, *m/z* 107.1 $[M + H - Gua]^+$. $^{15}N_5$-N7G-NOR-Cl was prepared analogously starting with $^{15}N_5$-2'-deoxyguanosine, and its structure was confirmed by UV and mass spectrometry. UV: λ_{max} 246 nm, λ_{min} 275 nm (pH 4.9); ESI^+-MS/MS: *m/z* 262.7 $[M + H]^+$, *m/z* 183.1 $[M + H - NH(CH_2)_2Cl]^+$, *m/z* 107.1 $[M + H - Gua]^+$.

N-[2-[cysteinyl]ethyl]-N-[2-(guan-7-yl)ethyl]amine (Cys-NOR-N7G)

Boc-Cys-OH (167 mg, 0.7541 mmol) was combined with N7G-NOR-Cl (335 mg, 0.9438 mmol) in 7 mL DMSO, and the reaction mixture was stirred at 37 °C for 72 h. The insoluble product was isolated by filtration and separated by semi-preparative HPLC on a Supelcosil LC-18-DB column (25 cm x 10 mm, 5 μm) eluted with a linear gradient of acetonitrile (B) in 20 mM ammonium acetate, pH 4.9 (A). Solvent composition was

changed linearly from 0 to 24% B in 24 min and further to 60% in 6 min. Under these conditions, Boc-Cys-NOR-N7G eluted as a sharp peak at 20.9 min. ESI^+-MS/MS: *m/z* 443.5 $[M + H]^+ \rightarrow$ *m/z* 343.5 $[M + H - Boc]^+$ and *m/z* 193.3 $[M + H - Boc - Gua]^+$. The Boc group was removed by incubating Boc- Cys-NOR-N7G with 50% TFA at room temperature for 45 min, and the resulting deprotected product (Cys-NOR-N7G) was purified using the same HPLC method (retention time = 10.4 min). Cys-NOR-N7G: ESI^+-MS/MS: *m/z* 342.1 $[M + H]^+ \rightarrow$ *m/z* 191.1 $[M + H - Gua]^+$. Cys-NOR-[$^{15}N_5$]-N7G was synthesized analogously starting with $^{15}N_5$7G-NOR-Cl. ESI^+-MS/MS: *m/z* 347.1 $[M + H]^+ \rightarrow$ *m/z* 191.1 $[M + H - Gua]^+$.

Cell Culture

Human fibrosarcoma (HT1080) cells were obtained from the American Type Culture Collection (Camden, NJ). Human Xeroderma Pigmentosum Complementation Group A (XPA, XPA deficient), human Xeroderma Pigmentosum Complementation Group D (XPD, XPD deficient), and XPD-derivative cells stably expressing XPD (XPD corrected) cells were obtained from NIGMS Human Genetic Cell Repository (Camden, NJ). Prof. Alexandra Sobeck (University of Minnesota) kindly provided FA-D2 cells (PD20, FANCD2 deficient), and FA-D2-derivative cells stably expressing FANCD2 (PD20 + FANCD2). The cells were maintained as exponentially growing monolayer cultures in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), in a humidified incubator at 37 °C with 5% CO_2.

Cytotoxicity Experiments

HT1080, XPA, PD20, and PD20 corrected cells were plated in Dulbecco's modified Eagle's medium containing 10% FBS at a density of 3 $\times 10^4$ cells/dish and permitted to adhere for 24 hours. Cells (1.0 x 10^5, in triplicate) were treated with 0 – 2000 μM PM, mechlorethamine, or DEB for 3 h at 37 °C. Following treatment, cell media was replaced, and the cells were grown for an additional 48 h at 37 °C. Cell viability was determined using an Alamar Blue assay[441] using a Synergy HI Microplate reader (BioTek, Winooski, VT).

Cell Treatment with PM and Isolation of DNA-Protein Crosslinks

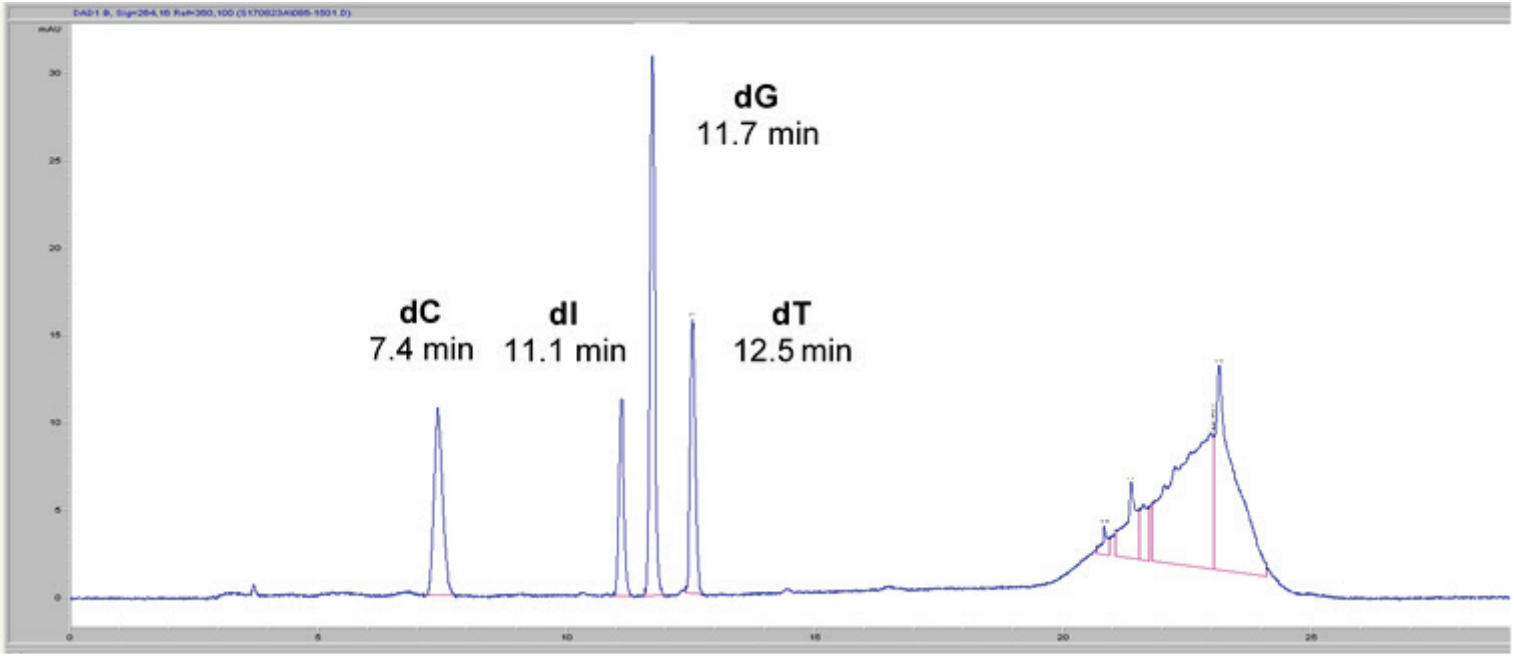
HT1080 cells in culture (~ 1 x $10^{7,}$ in triplicate) were treated with increasing concentrations of PM (0, 50, 100, 250, and 500 μM) for 3 h at 37 °C. Following treatment, the cells were washed with phosphate-buffered saline (PBS) and re-suspended in 3 mL PBS buffer. DPC-containing DNA was isolated by a modified phenol-chloroform procedure as described previously.[287, 440] In brief, the cells were lysed by adding an equal volume of 2X cell lysis buffer (20 mM Tris-HCl/10 mM $MgCl_2$/2% Triton-X100/0.65 M Sucrose), incubated on ice for 5 min, and centrifuged at 2000 g for 10 min at 4 °C. The nuclear pellets were re-suspended in saline-EDTA solution (75 mM NaCl/24 mM EDTA/1% sodium dodecyl sulfate) containing RNase A (10 μg/mL) and a protease inhibitor cocktail (1 mM PMSF; 1 μg/mL pepstatin; 0.5 μg/mL leupeptin; 1.5 μg/mL aprotinin). The resulting nuclei were incubated for 2 h at 37 °C with gentle shaking. To the nuclear lysates, an equal amount of tris-saturated phenol was added, and the solutions were mixed for 5 minutes. Following centrifugation at 1500 RPM for 10 minutes, the aqueous layer containing DNA and the DPC-containing interface was collected. This material was

re-extracted with phenol/chloroform as described above two additional times, and DPC-containing DNA was precipitated with cold isopropanol. DNA amounts and its purity were estimated using UV spectrophotometry (A_{260}) and subsequently determined by quantitation of dG in enzymatic hydrolysates as described below. Typical DNA yields from 10 million cells were 40-60 μg.

DNA Quantitation by dG Analysis

To quantify DPC-containing DNA extracted from cells and to detect any potential RNA contamination, 5 μg aliquots of DNA were taken and digested to 2'-deoxynucleosides in the presence of phosphodiesterase I (120 mU), phosphodiesterase II (105 mU), DNase (35 U) and alkaline phosphatase (22 U) in 20 μL 10 mM Tris-HCl/15 mM $ZnCl_2$ (pH 7.0) for 18 h at 37 °C. Quantitative analysis of dG in enzymatic digests was conducted by HPLC-UV on an Agilent Technologies 1100 HPLC system equipped with a diode array UV detector and an autosampler. Samples were loaded onto an Atlantis T3 C18 column (2.1 × 150 mm, 5 μm, from Waters Corporation, Milford, MA) and eluted with a gradient of 5 mM ammonium formate, pH 4.0 (A) and methanol (B). Solvent composition was changed linearly from 3 to 30% B over 15 min, increased further to 80% B over 3 min, held at 80% B for 1 min, and returned to 3% B over 2 min, where it was kept for the final 8 min of the HPLC run. UV absorbance was monitored at 260 nm. With this method, dG eluted as a sharp peak at 11.7 min (**Figure 2-1**). dG amounts were determined by comparing HPLC peak areas to a calibration curve constructed by injecting known dG amounts.

Figure 2-1: Representative trace of DNA dG analysis.



Mass Spectrometric Identification of Cross-Linked Proteins

To identify cellular proteins that become covalently trapped on DNA following exposure to PM, HT1080 cells (1 x 10^7, in triplicate) were treated with 100 μM PM for 3 h at 37 °C. Chromosomal DNA containing any covalently attached proteins was isolated by the modified phenol/chloroform extraction method described above.[287, 440] DPC-containing DNA (30 μg) was subjected to neutral thermal hydrolysis to release protein-guanine conjugates (**Scheme 2-2**), dried under vacuum, and reconstituted in 1 × NuPAGE Sample Buffer (Invitrogen, Carlsbad, CA). Proteins were resolved by NuPAGE Novex 12% Bis-Tris Gels (Invitrogen, Carlsbad, CA) and stained with SimplyBlue Safe stain (Invitrogen, Carlsbad, CA).

Gel lanes were divided into five sections encompassing the molecular weight range of 5 – 250 kDa. Each section of the gel (260 – 110 kDa, 110 – 80 kDa, 80 – 50 kDa, 50 – 30 kDa, 30 – 5 kDa) was excised and further diced into 1 mm pieces. The proteins present within the gel pieces were subjected to in-gel tryptic digestion by standard methods. In brief, gel pieces were washed with 25 mM ammonium bicarbonate, protein disulfide linkages were reduced with DTT (300 mM), and the resulting thiols were alkylated with saturated iodoacetamide. Gel pieces were dehydrated by incubation with acetonitrile, dried under vacuum, and reconstituted in 25 mM ammonium bicarbonate buffer (pH 9.0). Mass spectrometry grade trypsin (1 μg) was added, and the samples were digested for 18 hours at 37 °C. The resulting tryptic peptides were extracted with 60% acetonitrile containing 0.1% aqueous formic acid, evaporated to dryness, and desalted using C18 ZipTips (Millipore, Temecula, CA). Samples were reconstituted in 0.1% formic acid for HPLC−ESI^+−MS/MS analysis.

HPLC-ESI$^+$-MS/MS analyses of tryptic peptides were conducted using an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific, Waltham, MA) equipped with an Eksigent nanoLC 2D HPLC pump, a nanospray source, and Xcalibur 2.1.0 software for instrument control. Peptide mixtures were loaded onto a Symmetry C18 trapping column (180 μm ×20 mm, Waters, Milford, MA) using 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B) at a solvent composition of 95% A and 5% B at 5 μL/min for 3 min. Typical injection volume was 8 μL. Following trapping, the HPLC flow was decreased to 0.3 μL/min and reversed to elute the peptides. The peptides were separated on a nanoHPLC column (75 μm i.d.,10 cm packed bed, 15 μm orifice) created by hand packing a commercially purchased fused-silica emitter (New Objective, Woburn MA) with Zorbax SB-C18 5 μm separation media (Agilent, Santa Clara, CA). The gradient program started at 5% B, followed by a linear increase to 60% B over 60 min and further to 95% B in 5 min. Liquid chromatography was carried out at an ambient temperature. The mass spectrometer was calibrated prior to each analysis, and the spray voltage was adjusted to ensure a stable spray. Typically, the MS tune parameters were as follows: spray voltage of 1.6 kV, a capillary temperature of 275 °C, and an S-lens RF level of 50%. MS/MS spectra were collected using data-dependent scanning, in which one full scan mass spectrum acquired in the Orbitrap detector (R = 30,000) was followed by eight MS/MS spectra acquired in the Orbitrap detector (R = 7,500) with an isolation width of 2.5 m/z, activation time of 30 ms, activation Q of 0.25, 35% normalized CID collision energy, 1 microscan, with an AGC setting of 2 x 10^5 and a maximum injection time of 100 ms. Dynamic exclusion was enabled for 60 s, and singly charged species were excluded from MS/MS analysis.

Spectral data were analyzed using an in-house developed software pipeline "TINT" that linked raw data extraction, database searching, and probability scoring. Raw data were extracted and converted to the mzXML format using ReadW. Spectra that contained fewer than 6 peaks or had measured total ion current (TIC) < 20 were excluded. Data was processed using the SEQUEST v.27 algorithm[442] on a high speed, multiprocessor Linux cluster in the Minnesota Supercomputing Institute at the University of Minnesota. Peptide spectra were searched against the UniProt Human Protein Database. Cysteine carboxamidomethylation (+57.0215 Da) was set as a fixed modification, and methionine oxidation (+15.9949 Da) was selected as a variable modification. Precursor mass tolerance was set to 10 ppm within the calculated mass, and fragment ion mass tolerance was set to 10 mmu of their monoisotopic mass. The identified peptides were filtered using Scaffold 3 software (Proteome Software, INC., Portland, OR), to a target false discovery rate (FDR) of 5%. The FDR was calculated with the following expression: FDR = (2R)/(R + F) × 100, where R is the number of passing reversed peptide identifications and F is the number of passing forward (normal orientation) peptide identifications. The second round of filtering removed proteins supported by less than two unique peptide fragments in the analyses. Any proteins also found in control (untreated) samples were excluded from the final list.

Western Blot Analysis

HT1080 cells (~1 x 10^7) were treated with NOR (0, 50, 100, 250, or 500 μM) for 3 h at 37 °C. Chromosomal DNA, along with any covalent DPCs, was extracted and quantified as described above. From each sample, 50 μg of DNA was subjected to neutral thermal hydrolysis (1 h at 70 °C) to release protein-guanine conjugates from the DNA

backbone. The proteins were resolved by NuPAGE Novex 12% Bis-Tris gels (Invitrogen, Carlsbad, CA) and transferred to Invitrolon PVDF filter paper membranes (0.45 μm pore size, Life Technologies, Carlsbad, CA). The membranes were blocked for 4 h in Tris-buffered saline-Tween 20 (TBST) containing 5% (w/v) bovine serum albumin. Following blocking, the membranes were incubated with the primary antibodies against vimentin, nucleophosmin, prohibitin-2, matrin-3, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), poly-(ADP-ribose) polymerase 1 (PARP), and histone-H4 overnight at 4 °C. The membranes were rinsed with TBST and incubated with the corresponding alkaline phosphatase-conjugated antibody for 1 h at room temperature. The gels were developed using SIGMA Fast BCIP/NBT (Sigma, St. Louis, MO) according to manufacturer's protocol.

Quantitation of Cys-NOR-N7G in NOR-Exposed Cells

HT1080 cells (1 x 10^7) were treated with increasing concentrations of NOR (0, 50, 100, 250, and 500 μM) for 3 h at 37 °C. Following treatment, the cells were washed with PBS, re-suspended in cell lysis solution (2 mL, Qiagen, Valencia, CA) containing 10 μg RNAse A, and incubated for 16 h at room temperature with gentle inversion. Proteinase K (10 μg) was added, and the samples were incubated for an additional 16 h at room temperature with gentle inversion. The lysates were mixed with 700 μL of protein precipitation solution (Qiagen, Valencia, CA), and the samples were vortexed for one minute. Following centrifugation at 1500 RPM for 10 minutes, the supernatant was collected, and DPC-containing genomic DNA was precipitated with cold isopropanol. DNA was quantified by dG analysis as described above.

DNA (100 μg) was subjected to neutral thermal hydrolysis (1 h at 70 °C) to release protein−guanine conjugates from the DNA backbone. Proteins were cleaved with trypsin (10 μg protein in 25 mM ammonium bicarbonate, overnight at 37 °C), and the resulting peptides were further digested to amino acids in the presence of proteinase K (10 μg in 100 μL water, overnight at 37 °C). The digests were spiked with Cys-[$^{15}N_5$]-NOR-N7G (internal standard for mass spectrometry, 200 fmol), followed by offline HPLC purification as follows. An Agilent Technologies HPLC system (1100 model) was used incorporating a diode array detector, an autosampler, and a fraction collector. A Synergi 4μ Hydro RP (4.6 × 250 mm, 5 μm) column (Sigma-Aldrich, St. Louis, MO) was eluted at a flow rate of 1 mL/min using a gradient of 15 mM ammonium formate, pH 4.9 (A) and acetonitrile (B). Fractions containing Cys-NOR-Gua and its ^{15}N-labeled internal standard (9.7 – 11.2 min) were collected, dried under vacuum, and reconstituted in 0.1% formic acid (20 μL).

Quantitative analyses of Cys-NOR-N7G were conducted with a Dionex UltiMate 3000 RSLC nanoHPLC system (Thermo Scientific, Waltham, MA) interfaced to a TSQ Vantage mass spectrometer (Thermo Scientific, Waltham, MA). Chromatographic separation was accomplished with a Hypercarb HPLC column (100 mm × 0.3 mm, 3 μm) eluted with a gradient of 0.1% formic acid (A) and acetonitrile (B) at a flow rate of 12 μL/min. Solvent composition was linearly changed from 4% to 30% in 13 min and further to 90% 1 min, kept at 90% for 1 min, then brought back to 4% in 1 min. Under these conditions, Cys-NOR-N7G and its internal standard (Cys-NOR-[$^{15}N_5$] N7G) eluted at ~ 9.6 min. Electrospray ionization was achieved at a spray voltage of 3200 V and a capillary temperature of 270 °C. Collision induced dissociation was performed with Ar as a collision

gas (1.5 mTorr) at a collision energy of 23 V. Instrument parameters were optimized for maximum response during infusion of a standard solution of Cys-NOR-N7G.

HPLC-ESI$^+$-MS/MS analysis of Cys-NOR-N7G was performed in the selected reaction monitoring mode by following the neutral loss of guanine from protonated molecules of the analyte (*m/z* 342.1 $[M + H]^+ \rightarrow$ 191.0 $[M + H - Gua]^+$) and the corresponding mass transitions corresponding to $^{15}N_5$- labeled internal standard (*m/z* 347.1 $[M + H]^+ \rightarrow$ 191.0 $[M + H - {}^{15}N_5 - Gua]^+$). Analyte concentrations were determined using the relative response ratios calculated from HPLC-ESI$^+$-MS/MS peak areas in extracted ion chromatograms corresponding to Cys-NOR-N7G and its internal standard.

Cys-NOR-N7G Quantitation in Wild Type and DNA Repair-Deficient Fibroblasts

Human fibrosarcoma cells (HT1080), NER-deficient Xeroderma pigmentosum cells (XPA), FANCD2-deficient cells (PD20), and the corresponding FANCD2-corrected cells (PD20-corrected, $1x10^7$) were treated with 250 μM NOR for 3 h at 37 °C. Following treatment, cell media was replaced, and the cells were allowed to recover at 37 °C for either 4 h (n = 6) or 0, 4, 12, or 24 h (n = 3) to allow for repair. DNA was isolated using Qiagen Cell Lysis Solution extraction procedure and quantified by dG analysis as described above. DNA samples (100 μg) were subjected to neutral thermal hydrolysis (1 h at 70 °C) to release protein-guanine conjugates. Conjugated peptides were further digested to amino acids in the presence of trypsin (10 μg protein in 25 mM ammonium bicarbonate, overnight at 37 ° C), followed by proteinase K (10 μg in 250 μL water, overnight at 37 °C). The digests were spiked with Cys-NOR-[$^{15}N_5$]-N7G internal standard (200 fmol), followed by

off-line HPLC purification as described above. Cys-NOR-N7G conjugates were quantified by isotope dilution HPLC-ESI$^+$-MS/MS as described above.

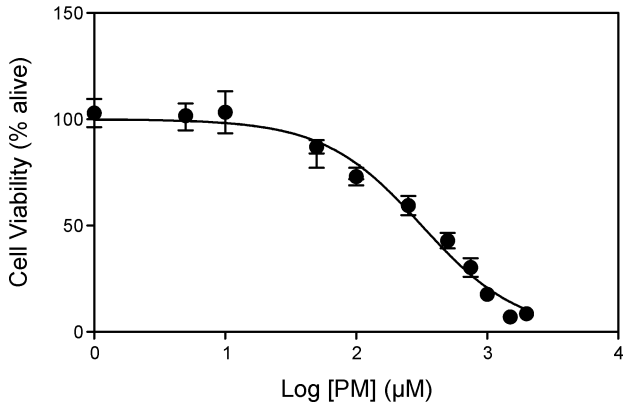
2.3. Results

2.3.1. Cytotoxicity experiments in human cell culture

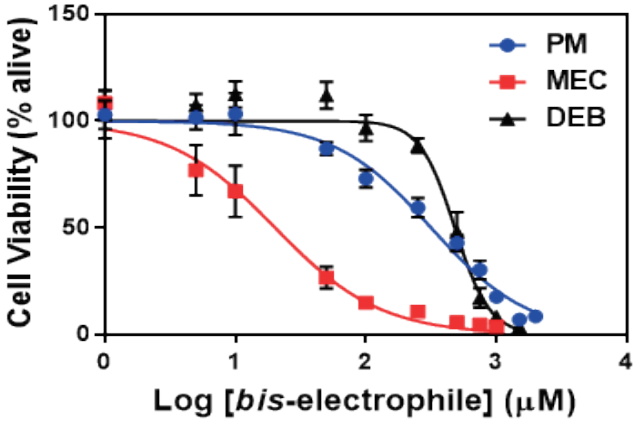
To characterize the toxicity of PM in human fibrosarcoma cell cultures, HT1080 cells (in triplicate) were treated with increasing amounts of PM, and their viability was determined by the Alamar Blue assay.[441] The fraction of live cells following treatment with 50, 100, 250, 500, 750, 1000, 1500, and 2000 μM PM was 87.0 ± 2.63, 73.0 ± 3.44, 59.5 ± 3.70, 42.9 ± 3.03, 30.3 ± 3.60, 17.7 ± 1.66, 7.08 ± 0.34, and 8.5 ± 0.33 % (**Figure 2-2A**). We found that PM was significantly less toxic than mechlorethamine, a structurally related nitrogen mustard investigated in our earlier publication (IC_{50} = 317 ± 42.9 and 20.2 ± 4.59, respectively),[287] but had similar toxicity to 1,2,3,4-diepoxybutane IC_{50} = 480.8 ± 41.9 (**Figure 2-2B**).

Figure 2-2: Cytoxicity of human fibrosarcoma (HT1080) cells to DPC-inducing agents. (**A**) HT1080 cytotoxicity (in triplicate) to 0 – 2000 μM PM. (**B**) Comparison of HT1080 cytotoxicity (in triplicate) to PM, 1,2,3,4-diepoxybutane, and mechlorethamine.

A.



B.



2.3.2. Concentration-dependent formation of DPCs in PM-treated human cells

To monitor PM-mediated DPC formation, human fibrosarcoma cells (HT1080) were treated with increasing concentrations of PM (0, 50, 100, 250, and 500 μM) for 3 h. DNA was extracted under mild conditions that preserved covalent DNA-protein conjugates. The samples were subjected to thermal hydrolysis to release proteins in the form of protein-guanine conjugates, which were resolved by gel electrophoresis (**Scheme 2-2** and **Figure 2-3A**). Although small amounts of endogenous DPCs were present in untreated cells (Lane 3 in **Figure 2-3A**), the amounts of cross-linked proteins increased in a concentration dependent manner (Lanes 5, 7, 9, 11 in **Figure 2-3A**, **Figure 2-3B**). In our previous studies, a 20-fold greater concentration of 1,2,3,4-diepoxybutane (DEB, 2 mM) was required to produce similar numbers of DPCs in HT1080 cells,[440] while comparable cross-linking efficiency was observed upon treatment with mechlorethamine.[287] This suggests that PM and NOR are equally efficient at forming DPCs, while DEB is less efficient than the two nitrogen mustards.

Scheme 2-2: Experimental strategy for isolation of DNA-protein cross-links from phosphoramide mustard-treated cells and their characterization by mass spectrometry.

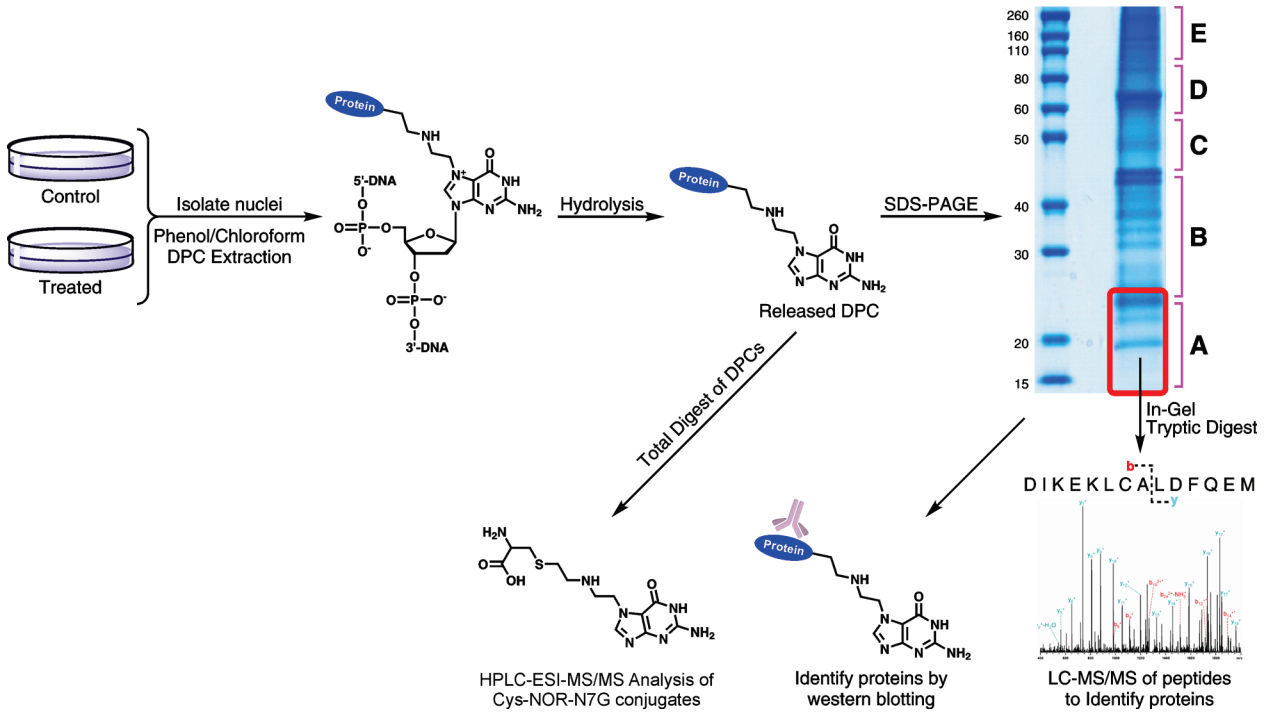
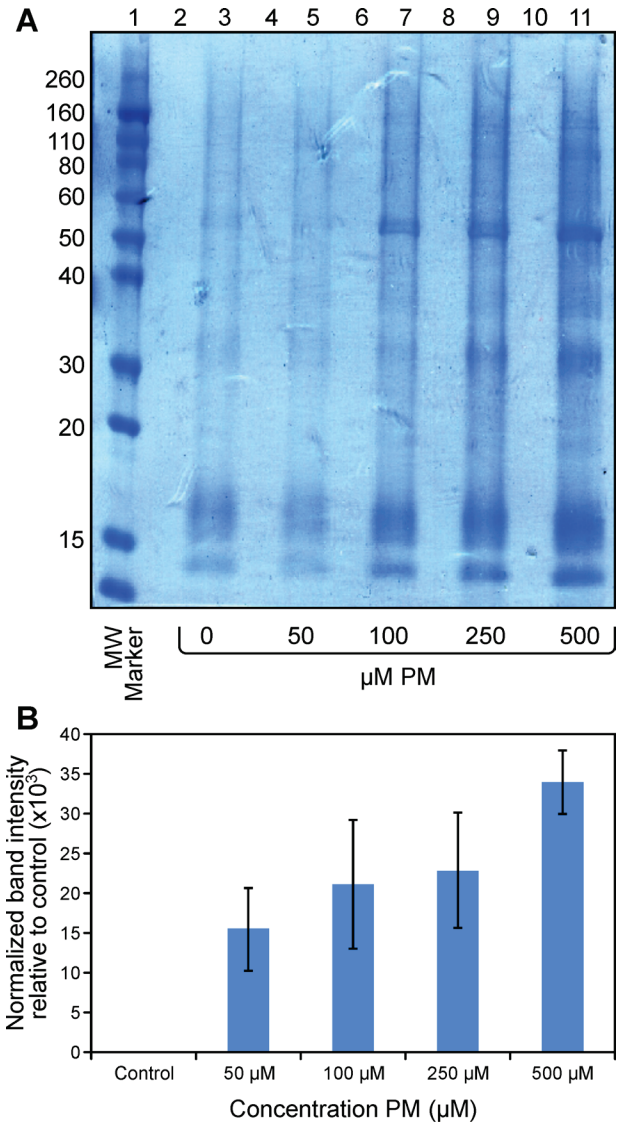


Figure 2-3: (**A**) Concentration-dependent formation of DPCs in HT1080 cells treated with 0 – 500 µM PM for 3 h. Lane 1: molecular weight marker. Lanes 2, 4, 6, 8, 10 are blank. Lane 3: 50 µM PM. Lane 5: 100 µM PM. Lane 7: 250 µM PM. Lane 9: 500 µM PM. (**B**) Densitometry measurements of the band intensity form Figure 2A, normalized to the untreated control.

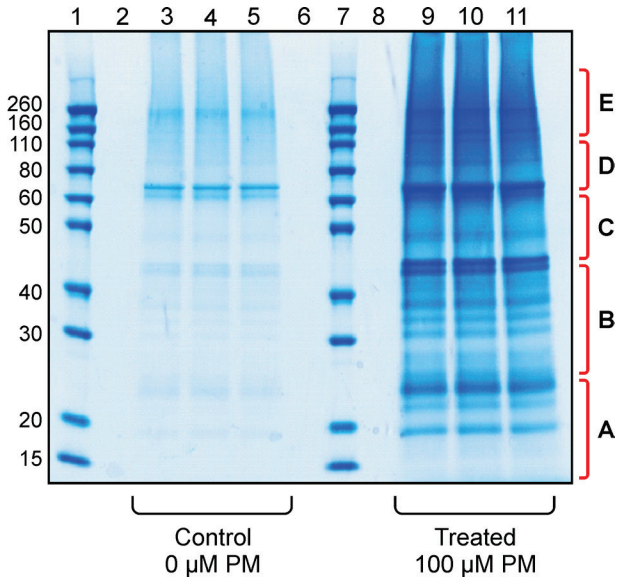


2.3.3. Identification of cross-linked proteins by mass spectrometry-based proteomics

To identify the proteins participating in DPC formation following exposure to PM, HT1080 cells (in triplicate) were treated with 100 μM PM for 3 h. Control cells were incubated with PM-free cell culture media. DPC-containing DNA was extracted as described above, and equal amounts of DNA from each sample (30 μg) were subjected to neutral thermal hydrolysis to release protein-guanine conjugates (**Scheme 2-2**), which were detected by gel electrophoresis.

SDS-PAGE analysis of protein-guanine conjugates isolated from PM-treated samples (Lanes 9-11 in **Figure 2-4**) revealed intense protein bands, while control samples contained considerably weaker bands corresponding to endogenous DPCs (Lanes 3 – 5 in **Figure 2-4**). The gel lanes from both experiments were divided into five sections covering the molecular weight range of 5 – 260 kDa (A – E in **Figure 2-4**) and cut out of the gel. Individual gel pieces of each region were subjected to in-gel tryptic digestion, and the resulting peptides were extracted from the gel, desalted, and subjected to HPLC-ESI$^+$-MS/MS analysis for protein identification.

Figure 2-4: DNA-protein cross-linking in human fibroblasts used for proteomics experiments. HT1080 cells (~5.0 x 10^6, in triplicate) were treated with 100 µM PM (lanes 9 – 11) or buffer only (lanes 3 – 5) for 3 h. Following DPC extraction and thermal hydrolysis to release guanine-protein conjugates, the cross-linked proteins were resolved by 12% SDS-PAGE and visualized by staining with SimplyBlue SafeStain. Proteins within the molecular weight ranges of 30 – 5 kDa (A), 50 – 30 kDa (B), 80 – 50 kDa (C), 110 – 80 kDa (D), and 260 – 110 kD (E) were excised from the gels, subjected to in-gel trypsin digestion, and analyzed by HPLC-ESI^+-MS/MS.



MS/MS analysis of tryptic peptides from cross-linked proteins yielded characteristic b and y series ions, which were used to determine their amino acid sequences and to identify the corresponding proteins. Protein identification was based on the sequence of at least two unique peptides. Any proteins that were also detected in control samples or were not present in all three treated samples were omitted from the final list. Using these criteria, a total of 134 proteins were identified (**Table 2-1**).

Proteins found to form covalent cross-links to chromosomal DNA following exposure to PM (**Table 2-1**) were categorized based on their cellular distribution, molecular function, and biological process using the GO database available from the European Bioinformatics Institute (http://www.ebi.ac.uk/GOA).[443] A large portion of the cross-linked proteins are classified as nuclear (67 total, 50%), including histones, high mobility group proteins, protein MAK16 homolog, and U1 small nuclear ribonucleoprotein 70 kDa (**Figure 2-5**, **Table 2-1**). This is not unexpected as nuclear proteins are most likely to form DPCs due to their proximity to nuclear DNA. The remaining proteins were classified as cytoplasmic (38 total, 28.4%), membrane bound/cytoskeleton (18 total, 13.4%), endoplasmic reticulum (7 total, 5.2%), mitochondrial (5 total, 3.7%), golgi apparatus (2, 1.5%) and unknown (2 total, 1.5%) (**Figure 2-5**). It is important to note that many of the proteins may be present in multiple cellular compartments due to their participation in many diverse biological processes. For example, S100 A9 is classified as a cytoplasmic protein, but it has been shown to enter the nucleus and bind the promoter region of complement component 4 (C3), upregulating C3 expression.[444]

Table 2-1: Proteins forming covalent cross-links to chromosomal DNA following exposure to PM.

Swiss-Prot ID	Identified Protein	% Coverage	No. of Unique Peptides	Total Spectra	Primary Biological Function	Protein MW (Da)
P53999	Activated RNA polymerase II transcriptional coactivator p15	37	4	4	*Transcriptional Regulation*	14395.9
O15347	High mobility group protein B3	19	3	3		22980.7
P39687	Acidic leucine-rich nuclear phosphoprotein 32 family member A	21	5	7		28586.1
Q99623	Prohibitin-2	10	2	2		33297.6
P16989	DNA-binding protein A	8	2	2		40089.4
Q14320	Protein FAM50A	8	2	2		40242.8
Q5BKZ1	Zinc finger protein 326	11	4	4		65653.5
P11142	Heat shock cognate 71 kDa protein	8	3	3		70899.8
Q8WVC0	RNA polymerase-associated protein LEO1	4	2	2		75405.3
Q9Y3T9	Nucleolar complex protein 2	3	2	2		84907.1
P43243	Matrin-3	13	10	11		94626.7
P55884	Eukaryotic translation initiation factor 3 subunit B	22	15	18		92482.8
Q8TDD1	ATP-dependent RNA helicase DDX54	2	2	2		98597.9
Q9BXP5	Serrate RNA effector molecule homolog	3	3	3		100669.7
Q9Y5B9	FACT complex subunit SPT16	16	15	17		119917.4
Q03701	CCAAT/enhancer binding protein zeta	2	2	2		120992.2

Swiss-Prot ID	Identified Protein	% Coverage	No. of Unique Peptides	Total Spectra	Primary Biological Function	Protein MW (Da)
Q9BQG0	Myb-binding Protein 1A	2	2	2		148858.3
P51532	Transcription activator BRG1	2	3	3		184649.6
O60216	Double-strand-break repair protein rad21 homolog	6	3	3	*DNA Repair*	71691.6
P12956	X-ray repair cross-complementing protein 5	22	10	12		69846.4
P84103	Splicing factor, arginine/serine-rich 3	30	5	5	*RNA Splicing/Processing*	19330.0
Q07955	Splicing factor, arginine/serine-rich 1	12	3	3		27745.1
Q32P51	Heterogeneous nuclear ribonucleoprotein A1-like 2	15	4	5		34225.4
P22626	Heterogeneous nuclear ribonucleoprotein A2/B1	10	3	3		37430.3
P08621	U1 small nuclear ribonucleoprotein 70 kDa	5	2	2		51558.4
O00566	U3 small nucleolar ribonucleoprotein protein MPP10	7	3	3		78866.8
Q15029	116 kDa U5 small nuclear ribonucleoprotein component	4	3	3		109438.1
Q15393	Splicing factor 3B subunit 3	2	2	2		135579.4
O75643	U5 small nuclear ribonucleoprotein 200 kDa helicase	2	4	4		244513.8

Swiss-Prot ID	Identified Protein	% Coverage	No. of Unique Peptides	Total Spectra	Primary Biological Function	Protein MW (Da)
P51665	26S proteasome non-ATPase regulatory subunit 7	18	3	3	*Apoptosis*	37025.6
P04083	26S proteasome non-ATPase regulatory subunit 3	19	9	9		60979.6
Q9NY61	Protein AATF	22	7	8		63135.0
P25205	DNA replication licensing factor MCM3	6	4	4	*DNA Replication*	90982.6
Q14566	DNA replication licensing factor MCM6	3	2	2		92890.2
P33778	Histone H2B type 1-B	20	3	14	*Chromatin Organization*	13950.8
Q92522	Histone H1x	10	2	2		22487.9
Q9P0M6	Core histone macro-H2A.2	7	2	3		40060.1
Q969G3	SWI/SNF-related matrix associated actin dependent regulator of chromatin subfamily E member 1	13	4	5		46649.7
Q14839	Chromodomain-helicase-DNA-binding protein 4	4	6	6		217994.6
Q12873	Chromodomain-helicase-DNA-binding protein 3	1	2	2		226597.2
P63173	60S ribosomal protein L38	29	2	4	*Translation*	8218.5
P62979	40S ribosomal protein S27a	40	2	3		9418.4
Q6NW1	Putative 60S ribosomal protein L13a-like MGC87657	29	3	3		12135.0
P83881	60S ribosomal protein L36a	30	4	4		12441.1

Swiss-Prot ID	Identified Protein	% Coverage	No. of Unique Peptides	Total Spectra	Primary Biological Function	Protein MW (Da)
Q5JNZ5	Putative 40S ribosomal protein S26-like 1	27	3	10		13002.2
P60866	40S ribosomal protein S20	19	2	3		13373.0
P62899	60S ribosomal protein L31	30	4	11		14463.2
P35268	60S ribosomal protein L22	19	2	3		14787.3
P62244	40S ribosomal protein S15a	31	4	4		14840.0
P62829	60S ribosomal protein L23	32	3	4		14865.9
P62847	40S ribosomal protein S24	29	4	9		15423.8
P08708	40S ribosomal protein S17	16	2	3		15550.5
P62266	40S ribosomal protein S23	15	2	3		15807.7
P62263	40S ribosomal protein S14	34	4	4		16272.9
P62249	40S ribosomal protein S16	28	4	6		16445.9
P46776	60S ribosomal protein L27a	20	3	3		16561.4
P62841	40S ribosomal protein S15	40	3	6		17040.8
P62277	40S ribosomal protein S13	36	6	7		17223.3
P30050	60S ribosomal protein L12	24	3	4		17819.1
P62280	40S ribosomal protein S11	35	5	5		18431.3
P46778	60S ribosomal protein L21	21	3	8		18565.0
P46783	40S ribosomal protein S10	15	2	4		18898.3
Q9NQ39	Putative 40S ribosomal protein S10-like	13	2	3		20120.6
Q02543	60S ribosomal protein L18a	21	3	4		20762.6
T1W3K1	60S ribosomal protein L17	41	6	11		21397.4
P32969	60S ribosomal protein L9	18	3	4		21863.7
P62081	40S ribosomal protein S7	47	6	10		22127.5
P46781	40S ribosomal protein S9	45	12	21		22592.5
P23396	40S ribosomal protein S3	27	5	5		26688.6
P62753	40S ribosomal protein S6	18	4	5		28681.7

Swiss-Prot ID	Identified Protein	% Coverage	No. of Unique Peptides	Total Spectra	Primary Biological Function	Protein MW (Da)
P18124	60S ribosomal protein L7	38	12	16		29227.7
P15880	40S ribosomal protein S2	24	6	6		31325.2
P05198	Eukaryotic translation initiation factor 2 subunit 1	10	2	2		36112.7
P20042	Eukaryotic translation initiation factor 2 subunit 2-like protein	16	5	5		37788.3
P60842	Eukaryotic initiation factor 4A-I	22	7	7		46155.3
O15371	Eukaryotic translation initiation factor 3 subunit D	11	5	5		63973.8
P55884	Eukaryotic translation initiation factor 3 subunit B	22	15	18		92482.8
Q14152	Eukaryotic translation initiation factor 3 subunit A	12	15	18		166574.9
Q04637	Eukaryotic translation initiation factor 4 gamma 1	9	13	14		175494.2
P05109	Protein S100 A8	32	3	4	*Inflammatory Response*	10835.0
P06702	Protein S100 A9	45	4	6		13242.3
P66905	Hemoglobin subunit alpha	22	2	3	*Oxygen Transport*	15257.6
Q13765	Nascent polypeptide-associated complex subunit alpha	26	4	5	*Protein Transport*	23383.3
Q96FZ7	Charged multivesicular body protein 6	10	2	2		23485.3
O76021	Ribosomal L1 domain-containing protein 1	11	4	4		54974.7
O00203	AP-3 complex subunit beta-1	5	4	4		121324.0
Q9P2E9	Ribosome-binding protein 1	14	16	21		152470.6
Q08378	Golgin subfamily A member 3	3	3	3		167356.0

Swiss-Prot ID	Identified Protein	% Coverage	No. of Unique Peptides	Total Spectra	Primary Biological Function	Protein MW (Da)
Q00610	Clathrin heavy chain 1	11	15	15		191618.9
Q13268	Dehydrogenase/reductase SDR family member 2, mitochondrial	11	3	3	*Cellular Homeostasis*	27438.4
Q15006	Tetratricopeptide repeat protein 35	18	4	4		34834.8
P00338	L-lactate dehydrogenase A chain	13	4	5		36689.2
P11201	78 kDa glucose-regulated protein	17	8	8		72334.7
P04083	Annexin A1	27	6	8		38715.9
Q96A33	Coiled-coil domain-containing protein 47	9	3	3		55874.9
P25705	ATP synthase subunit alpha, mitochondrial	12	6	6		59752.1
Q9NR30	Nucleolar RNA helicase 2	6	4	4		87346.0
P36776	Lon protease homolog, mitochondrial	6	5	6		106491.3
P63104	14-3-3 protein zeta/delta	17	3	4	*Cellular Signaling*	27745.9
P31946	14-3-3 protein beta/alpha	9	2	5		28083.1
P62258	14-3-3 protein epsilon	12	3	3		29175.0
Q9UBS4	DnaJ homolog subfamily B member 11	14	3	3		40514.6
Q3TIV5	Zinc finger CCCH domain-containing protein 15	10	3	3		48603.7
Q13283	Ras GTPase-activating protein-binding protein 1	9	3	4		52162.8
Q1KMD3	Heterogeneous nuclear ribonucleoprotein U-like protein 2	9	6	7		85105.2
P05556	Integrin beta-1	6	4	4		88415.1
O14786	Neuropilin-1	19	12	13		103136.6

Swiss-Prot ID	Identified Protein	% Coverage	No. of Unique Peptides	Total Spectra	Primary Biological Function	Protein MW (Da)
P08865	40S ribosomal protein SA	19	3	3	*Cell Structure/Architecture*	32854.1
P60709	Actin, cytoplasmic 1	19	5	6		41737.8
P68363	Tubulin alpha-1B chain	6	2	5		50151.7
P55081	Microfibrillar-associated protein 1	17	5	6		51958.7
P17661	Desmin	4	2	8		53536.6
P41219	Peripherin	4	2	10		53652.0
P08670	Vimentin	67	30	142		53652.7
Q16352	Alpha-internexin	6	3	4		55392.3
Q07065	Cytoskeleton-associated protein 4	48	23	27		66022.2
Q16643	Drebrin	6	2	2		71428.6
P23284	Peptidyl-prolyl cis-trans isomerase B	21	4	4	*Protein Folding*	23743.2
Q9UKD2	mRNA turnover protein homolog 4	16	3	3	*Ribosome Assembly*	27561.3
Q99848	Probable rRNA-processing protein EBP2	27	6	8		34852.9
Q15050	Ribosome biogenesis regulatory protein homolog	24	7	9		41194.4
O00541	Pescadillo homolog	6	4	4		68004.9
Q14692	Ribosome biogenesis protein BMS1	4	3	3		145812.2
P54652	Heat shock-related 70 kDa protein 2	5	3	3	*Cell Cycle*	70023.0
Q5TAP6	U3 small nucleolar RNA-associated protein 14 homolog A	12	7	7		87979.4
P46087	Putative ribosomal RNA methyltransferase NOP2	9	5	5		89303.6
Q8WTT2	Nucleolar complex protein 3 homolog	3	2	2		92552.2

Swiss-Prot ID	Identified Protein	% Coverage	No. of Unique Peptides	Total Spectra	Primary Biological Function	Protein MW (Da)
A0LTM5	Cell division protein 11B	3	2	2		92709.0
O75400	Pre-mRNA-processing factor 40 homolog A	5	5	5		108807.3
Q14980	Nuclear mitotic apparatus protein 1	5	9	10		238259.7
Q9BXY0	Protein MAK16 homolog	15	3	4	*Unknown*	35370.0
Q14257	Reticulocalbin-2	32	9	12		36877.6
Q9NZM5	Glioma tumor suppressor candidate region gene 2 protein	14	6	6		54390.7

With respect to their molecular function, the majority of the cross-linked proteins are known to bind DNA and RNA (69 total, 37.7%). In addition, 10 of the proteins are known to participate in transcription factor binding, while 7 are involved in translation initiation factor binding (**Figure 2-5**). With regard to biological processes, many of the cross-linked proteins are known to be involved in transcriptional regulation (20 total, 12.0%), mRNA/RNA processing (17 total, 10.2%), translation (40 total, 23.9%), DNA damage response (8 total, 4.8%), and chromatin organization (7, 3.8%) (**Figure 2-5**). Included in this group are the nucleolar complex protein 2 (inhibitor of histone deacetyltransferase and p53/TP53-regulated target promoters),[445] Myb-binding protein 1A (binds sequence-specific DNA-binding proteins),[446, 447] double-strand-break repair protein rad21homolog,[448] and chromodomain-helicase-DNA-binding protein 4 (regulation of homologous recombination DNA repair).[449]

Some of the proteins were identified as structural constituents of the ribosome (27 total, 14.8%) (**Figure 2-5**). These include the 40S ribosomal protein S6, 60S ribosomal protein L7, 40S ribosomal protein S2, and 40S ribosomal S3 (**Table 2-1**). This is not a result of RNA contamination in our experiments, because HPLC analysis of enzymatic digests has confirmed that DNA isolated by our methodology was free of RNA. An alternative explanation is that many of known RNA binding proteins may also have an affinity for DNA. We also observed proteins involved in cell signaling (15 total, 8.9%) such as the heterogeneous nuclear ribonucleoprotein U-like protein 2, integrin β-1,[450] and neurophilin 1,[451] cell cycle/homeostasis (21 total, 11.5%) such as the Lon protease homolog[452] and dehydrogenase/reductase SDR family member 2,[453] and cell

structure/architecture (11 total, 6.6%) such as actin,[454] tubulin alpha-1B chain, and microfibrillar associated protein 1 (**Table 1**).

Many of the gene products identified in our study were counted into multiple GO categories due to their diverse cellular functions and their involvement in multiple biological processes. For example, the FACT complex subunit SPT16 is involved in regulating transcription from RNA polymerase II promoters,[455-457] cellular response to DNA damage stimulus,[458] and DNA replication.[459] As a result, it is listed under three GO annotation categories: DNA repair, DNA replication, and transcriptional regulation. It is also possible that the GO annotation does not take into account every protein's secondary cellular distributions and molecular functions. The identities of a subset of proteins detected by mass spectrometry-based proteomics were further confirmed by western blot analysis using commercially available antibodies (**Figure 2-6**).

Figure 2-5: GO annotations for the proteins involved in PM-induced DPC formation in human HT1080 cells according to cellular distribution, molecular function, and biological processes. The numbers of proteins falling into each category are included.

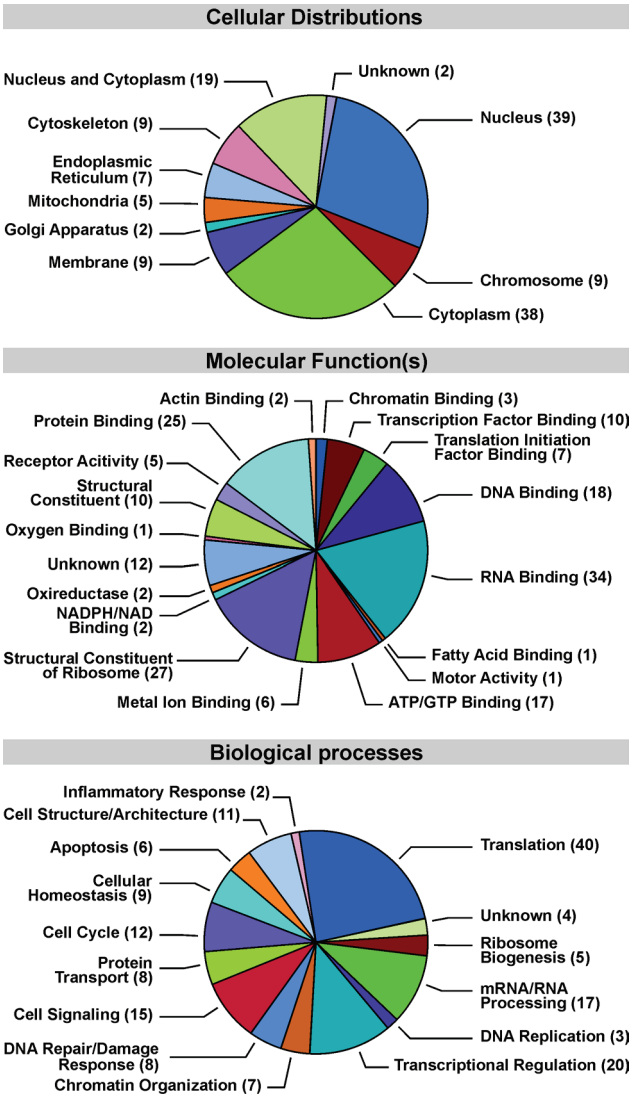
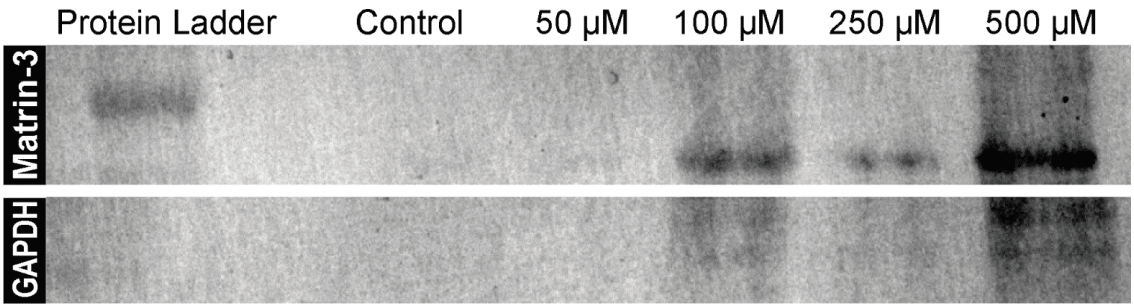


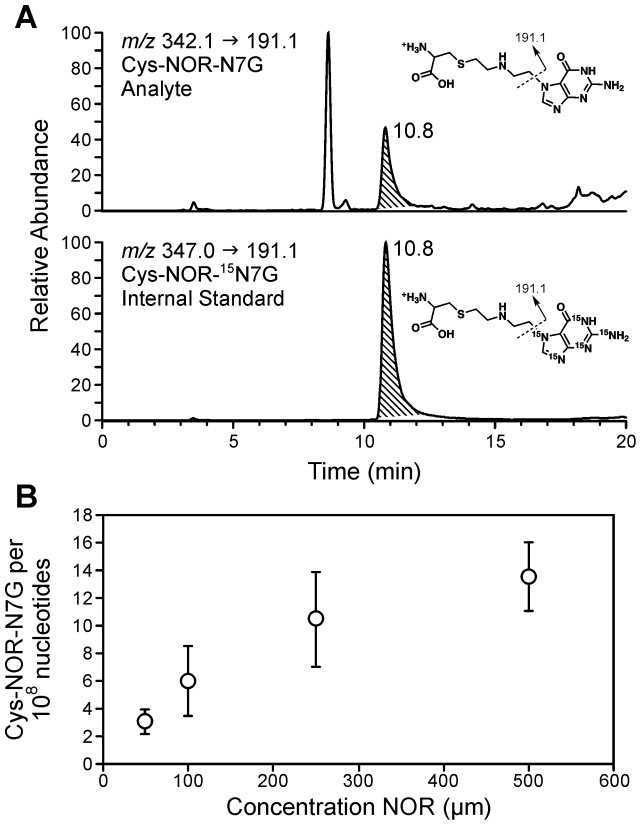
Figure 2-6: Western blot analysis of concentration-dependent DPC formation identifying Matrin-3 and GAPDH.



2.3.4. HPLC-ESI-MS/MS detection of Cys-NOR-N7G conjugates

To quantify DPC formation in NOR-treated fibroblasts, HT1080 cells (in triplicate) were treated with 50, 100, 250, or 500 μM NOR, and DPC-containing chromosomal DNA was extracted as described above. Following thermal hydrolysis to release DNA and proteolytic digestion of proteins to amino acids, isotope dilution HPLC-ESI$^+$-MS/MS was used to quantify N-[2-[cysteinyl]ethyl]-N-[2-(guan-7-yl)ethyl]methylamine (Cys-NOR-N7G) conjugates, which correspond to cross-linking between the N7-position of guanine in DNA and cysteine thiols of proteins (**Scheme 2-2** and **Figure 2-7A**). Cys-NOR-N7G adducts levels increased linearly with increasing concentrations of NOR, ranging between 3 and 14 adducts per 10^8 nucleotides (**Figure 2-7B**). These results confirm that PM and NOR treatment induces DPC formation between the N7-position of guanine and the cysteine thiols of proteins, as was previously reported for other nitrogen mustard drugs.[287, 435]

Figure 2-7: Representative HPLC-ESI$^+$-MS/MS trace (**A**) and quantitative results for Cys-NOR-N7G conjugates in HT1080 cells treated with 0-500 μM NOR for 3 h (**B**). Following treatment, genomic DNA containing DPCs was extracted, hydrolyzed to release DNA-protein conjugates, and digested by proteinase K to amino acids. NOR-induced Cys-NOR-N7G conjugates were analyzed by isotope dilution HPLC-ESI$^+$-MS/MS using the corresponding $^{15}N_5$-labeled internal standard.



2.3.5. Adduct formation in repair deficient cells

To gain insight into potential mechanisms of repair of NOR-induced DPCs in human cells, we examined PM toxicity and DPC formation in repair proficient and repair deficient cells. HT1080 (wild type), XPA (NER deficient), XPD (NER deficient), PD20 (FANCD2 deficient), XPD corrected cells (XPD proficient), and PD20 corrected cells (FANCD2 proficient) were employed in these experiments. Both NER-deficient Xeroderma pigmentosum cells (XPA and XPD) cells and Fanconi Anemia pathway-deficient cells (PD20) exhibited an increased sensitivity towards PM than the corresponding wild type cells (IC_{50} = 146 ± 31.2 and 168 ± 35.7 respectively, **Table 2-2**). PD20-corrected cells had similar sensitivity as HT1080 cells (IC_{50} = 313 ± 64.9 vs 317 ± 42.9, **Table 2-2**). These results suggest that both NER and FA repair pathways help protect cells against PM-mediated toxicity.

To determine whether the observed differences in sensitivity to PM can be attributed to impaired removal of DPC lesions, HT0180, XPA, PD20, and PD-corrected cells were treated with NOR (250 μM) and allowed to recover in drug-free media for 4h to allow for repair of NOR-induced DPCs. HPLC-ESI^+-MS/MS analyses have revealed that NER-deficient XPA and XPD cells had the highest numbers of Cys-NOR-N7G adducts (24.3 ± 3.5 adducts/10^8 nucleotides, **Figure 2-8A**) as compared to HT1080 (8.7 ± 1.5 adducts/10^8 nucleotides), PD20 (5.2 ± 1.3 adducts/10^8 nucleotides) and PD20-corrected cells (5.6 ± 1.6 adducts/10^8 nucleotides) (**Figure 2-8A**). To measure NOR-induced DPC repair over time, HT1080, XPA, PD20, and PD-corrected cells were treated with NOR (250 μM) and allowed to recover in drug-free media for 0, 4, 12, or 24 h. HPLC-ESI^+-MS/MS analysis revealed that XPA cells had the highest Cys-NOR-N7G adduct levels at

each time point measured (14.7 ± 6.2, 19.2 ± 5.1, 10.2 ± 6.1, and 7.2 ± 3.3 at 0, 4, 12, and 24 h respectively, **Figure 2-8B**). After 24 hours of repair, Cys-NOR-N7G adduct levels in all four cell types decreased to below 10 adducts/10^8 nucleotides (**Figure 2-8B**). Taken together, the results shown in **Table 2-2** and **Figure 2-8** are consistent with an involvement of nucleotide excision repair in the removal of cyclophosphamide-mediated DPCs. Fanconi Anemia pathway deficient cells (PD20) were more sensitive to PM (**Table 2-2**), but contained similar Cys-NOR-N7G adducts levels as compared to corrected cells (**Figure 2-8**). This could be rationalized by FA-mediated removal of other toxic DNA lesions induced by PM, such as DNA-DNA cross-links.[42-44]

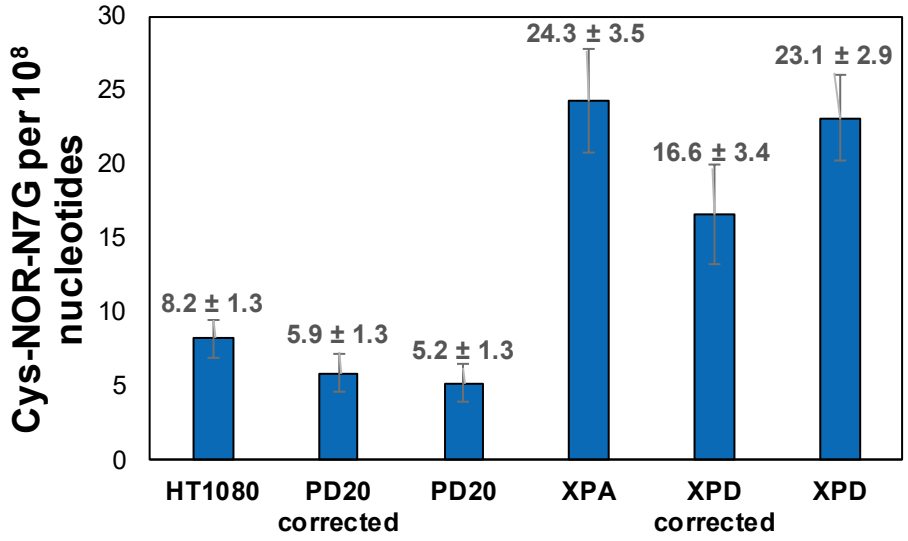
Table 2-2: Cell cytotoxicity of HT1080 (wild type), XPA (NER-deficient), XPD (NER-deficient), XPD corrected (NER-proficient), PD20 (FANCD2-deficient), PD20 corrected (PD20 + FANCD2) to PM.

Cell Type	IC_{50} [μM
HT1080	317 ± 43
PD20 Corrected	313 ± 65
XPD Corrected	162 ± 34
PD20	168 ± 34
XPD	89 ± 17
XPA	146 ± 31

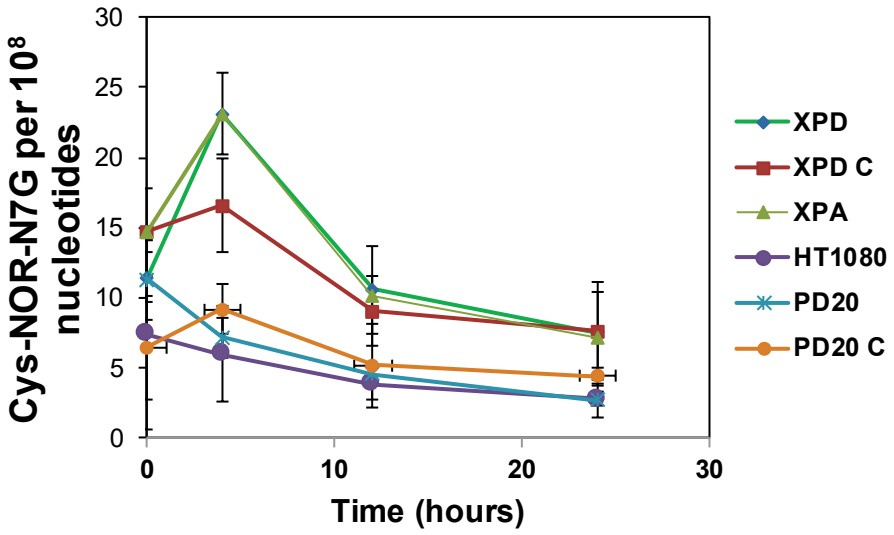
Cell Type	IC50 [μM
HT1080	317 ± 43
PD20 Corrected	313 ± 65
XPD Corrected	162 ± 34
PD20	168 ± 34
XPD	89 ± 17
XPA	146 ± 31

Figure 2-8: Quantitation of Cys-NOR-N7G formation and repair in cells. HT1080, XPA, XPD, XPD-corrected, PD20, and PD20- corrected cells treated with 250 μM NOR for 3 h. Following treatment, NOR-containing media was removed, and the cells were allowed to recover in fresh media for 4 h (**A**) or over a time frame of 0 – 24 h (**B**).

A.



B.



2.4. Discussion

DNA-protein cross-links are bulky DNA adducts that can form upon exposure to *bis*-electrophiles, reactive oxygen species, and transition metals.[262, 263] DPC formation is thought to contribute to the biological activity of *bis*-electrophiles. For example, the cytotoxicity and mutagenicity of dibromoethane and diepoxybutane is enhanced in bacteria overexpressing O^6-alkylguanine DNA transferase (AGT) protein due to the formation of covalent AGT-DNA conjugates.[284] Furthermore, proteins functionalized with DNA-reactive 2-hydroxy-3,4-epoxybutyl groups induced cell death and mutations in human cells.[336]

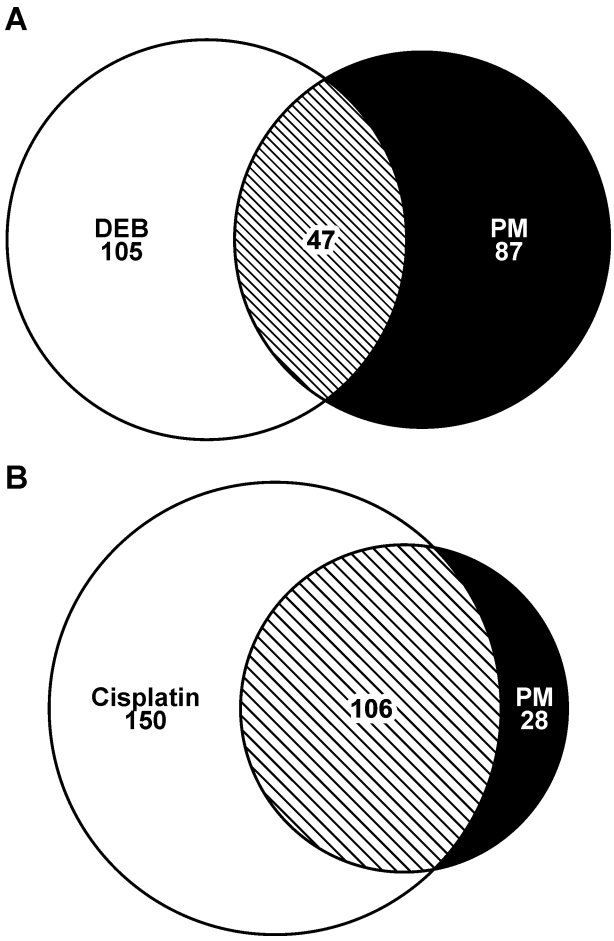
Identification of the proteins involved in DPC formation in living cells and determining the mechanisms of DPC repair are important for our understanding of the etiology of cardiovascular disease, age-related neurodegeneration, and cancer. DPCs accumulate in the heart and brain tissues of mice with age[270] and are generated as a result of ischemia-reperfusion injury (**Chapter 3**). DPC formation is likely to contribute to both

target and off-target toxicity of DNA *bis*-alkylating agents commonly used as antitumor drugs, and single nucleotide polymorphisms in DPC repair genes may contribute to inter-individual differences in response to antitumor therapy.

In the first part of this study, a mass spectrometry-based approach was used to characterize DNA-protein cross-linking following exposure of human cells to two biologically active metabolites of CPA (PM and NOR, **Scheme 2-1**). We found that PM and NOR induced covalent cross-links between DNA and over 130 cellular proteins, including those participating in chromatin remodeling, translation, DNA replication, DNA repair, RNA metabolism, transcriptional regulation, and apoptosis (**Table 2-1**). These proteins are involved in a variety of cellular functions including transcriptional regulation (e.g. prohibitin-2, protein FAM50A, and transcription activator BRG1), RNA splicing/processing (e.g. splicing factor, arginine/serine-rich 3, splicing factor 3B subunit 3, and 116 kDa U5 small nuclear ribonucleoprotein component), chromatin organization (e.g. core histone macro-H2A.2, SWI/SNF-related matrix associated actin dependent regulator of chromatin subfamily E member 1, chromodomain-helicase-DNA-binding protein 4), protein transport (e.g. charged multivesicular body protein 6, AP-3 complex subunit beta-1, and clathrin heavy chain1), cellular signaling (14-3-3 protein zeta/delta, DnaJ homolog subfamily B member 11, and neuropilin-1), and cell structure/architecture (actin, cytoplasmic-1, desmin, and vimentin).[460-462] When the list of proteins participating in cross-linking to DNA in the presence of PM was compared to an analogous lists proteins cross-linked to DNA in the presence of DEB,[440] a total of 47 proteins (30.9%) were found in common (**Figure 2-9A**). When compared to a list of cisplatin cross-linked proteins, 106 (41.4%) proteins were found in common (**Figure 2-9B**). These results suggest that nitrogen

mustards and platinum drugs may target the same group of proteins for cross-linking to DNA, while a distinct group of proteins is cross-linked by DEB. Further studies will establish whether these differences in specificity could contribute to different biological outcomes of DEB (known carcinogen)[212, 463, 464] and cyclophosphamide (useful antitumor agent).[326, 378, 439]

Figure 2-9: Venn diagrams showing the overlaps between the identified proteins that form DPCs in human cells following exposure to PM and DEB (**A**) and cisplatin (**B**).



HPLC-ESI$^+$-MS/MS analysis have revealed that similarly to other nitrogen mustards, PM and NOR cross-link cysteine thiols of proteins to the N7 position of guanine in DNA. Cys-NOR-N7G conjugates were formed in a concentration-dependent manner in

NOR-treated cells (**Figure 2-7**). Although N7-guanine cross links are expected to have limited hydrolytic stability, they may persist long enough to affect cellular replication and transcription, potentially leading to toxicity and mutations.

The biological outcomes of DNA-protein cross-linking are currently under active investigation. Nakano et al. have shown that covalent DPCs larger than 16.0 kDa blocked the progression of DNA helicases, thus preventing the unwinding of double stranded DNA ahead of the replication fork.[264] In our experiments with recombinant human DNA polymerases, we found that proteins and large peptides (> 20 amino acids) conjugated to the C-5 position of dT blocked DNA synthesis in the presence of lesion bypass polymerases η and κ, while smaller DNA-peptide conjugates (< 10-mer) were bypassed by in an error-prone manner, giving rise to deletions and point mutations.[265, 266] Recent *in vivo* studies of model peptide cross-links in human embryonic kidney cells have provided preliminary evidence for the ability of these conjugates to cause both targeted and off-target mutations (Pande et al., manuscript in preparation).

The processes by which covalent protein-DNA complexes are recognized as DNA damage and removed from genomic DNA are incompletely understood. Recently, Duxin *et al* used a plasmid containing a site-specific DPC to demonstrate that in *Xenopus* egg extracts, DPC repair is coupled to DNA replication.[430] These authors proposed that the collision of the replisome/CMG helicase with a DPC located on the DNA leading strand triggers proteolytic degradation of the protein constituent of DPC.[430] In the yeast, the metalloprotease Wss1 has been identified as the DPC-specific protease.[421] The resulting DNA-peptide cross-links may serve as substrates for nucleotide excision repair (NER).[309,

[426, 427] Alternatively, Nakano *et al* proposed that large DPCs do not undergo proteolysis, but rather are directly repaired by the HR pathway.[424, 425]

Our quantitative isotope dilution HPLC-ESI$^+$-MS/MS results are consistent with a role for NER in removal of NOR-induced DPCs (**Figure 2-8**). Cys-NOR-N7G numbers were significantly higher in cells deficient in nucleotide excision repair as compared with repair proficient cells. In contrast, the presence of functional FA pathway (PD 20 vs PD20 corrected) did not influence the numbers of Cys-NOR-N7G adducts observed in NOR-treated human cells (**Figure 2-8**). These results are consistent with a model that in replicating human cells, NOR-induced DPCs are proteolytically digested to smaller DNA-NOR-peptide conjugates, which are subject to repair by NER. Further investigations are needed to elucidate the relative contributions of NER and HR in protecting human cells form toxic effects of DPCs and to establish the role of DNA-specific proteases such as Dvc1/Spartan[421, 429] in initiating DPC repair.

2.5. Conclusions

In conclusion, a mass spectrometry-based proteomics approach was employed to characterize PM- and NOR-mediated DNA-protein cross-linking in human cells. Following treatment of human fibrosarcoma cells (HT1080) with cytotoxic concentrations of PM, over 130 proteins were found to be covalently trapped to DNA, including those involved in transcriptional regulation, RNA splicing/processing, chromatin organization, and protein transport. HPLC-ESI$^+$-MS/MS analysis of proteolytic digests of DPC-containing DNA from NOR-treated cells revealed a concentration-dependent formation of *N*-[2-[cysteinyl]ethyl]-N-[2-(guan-7-yl)ethyl]amine (Cys-NOR-N7G) conjugates,

confirming that it cross-links cysteine thiols of proteins to the N7 position of guanines in DNA. A sensitive and accurate Cys-NOR-N7G HPLC-ESI-MS/MS assay was developed to investigate the formation and repair of PM-induced DPCs. Cys-NOR-N7G adduct numbers were higher in NER-deficient Xeroderma pigmentosum cells (XPA and XPD) as compared with repair proficient cells, suggesting that NER is crucial for the repair of NER.

3. Oxidative Cross-Linking of Proteins to DNA Following Ischemia-Reperfusion Injury

3.1. Introduction

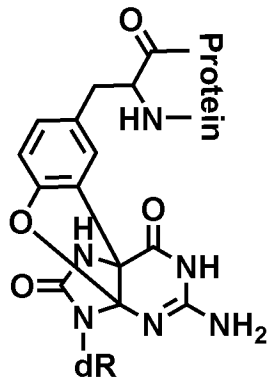
Heart disease is the leading cause of mortality for both men and women in the United States, accounting for 1 out of every 4 deaths.[364, 365] Coronary heart disease disrupts the flow of oxygenated blood to the heart, potentially leading to myocardial infarction (MI).[364, 365] Current MI therapies are based on the revascularization of the injured area using thrombolytic or angioplasty techniques in an attempt to rapidly return oxygenated blood flow to the damaged myocardium. Reperfusion-induced return of oxygenated blood to the affected myocardium, however, can lead to reactivation of the electron transport chain, producing large amounts of reactive oxygen species (ROS) such as superoxide anions ($O_2^{\bullet-}$), hydrogen peroxide (H_2O_2), and hydroxyl radicals ($OH^{\bullet}$).[465, 466] Furthermore, neutrophils accumulating in the infarcted myocardium can mediate tissue damage by releasing matrix-degrading enzymes and peroxides for downstream signaling.[370-372]

ROS generated during ischemia-reperfusion injury can cause oxidative damage to cellular biomolecules including lipids, proteins, and DNA.[367, 369] Several types of ROS-induced DNA damage can be formed including DNA strand breaks, nucleobase monoadducts, and covalent DNA-DNA and DNA-protein cross-links (DPCs).[262, 263, 359, 360] DPCs are super-bulky DNA lesions containing proteins irreversibly attached to DNA strands.[262, 263, 467] ROS-induced DPCs include covalent cross-links between ε-amino group of lysine and spirodiiminodihydantoin in DNA,[293] lysine and guanidinohydantoine in

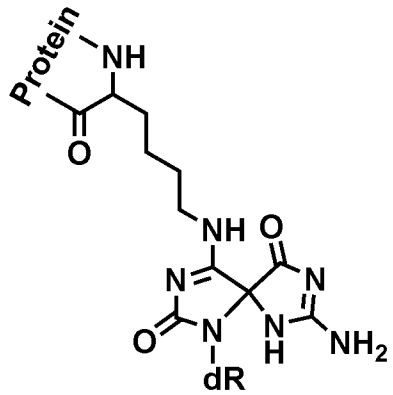
DNA,[294] tyrosine, and 8-oxoguanosine in DNA,[293] lysine and guanine in DNA (N^{ε}- lysine guanin-8-yl cross-links),[293, 294, 361] and tyrosine and thymidine in DNA (dT-Tyr) (**Scheme 3-1**).[362, 363] If not repaired, DPCs can induce DNA fragmentation and cell death.[262, 263, 467]

In the present study, we employed left anterior descending artery (LAD) ligation/reperfusion surgery [468] to examine the formation of radical-induced DPCs in cardiomyocytes after ischemia-reperfusion injury in a rat model. A sensitive isotope dilution nanoLC-ESI$^+$-MS/MS assay developed in our laboratory was used to quantify the formation of hydroxyl radical-mediated DPCs, while mass spectrometry based quantitative proteomics was employed to examine the proteins participating in DPC formation and to detect changes in the global cardiomyocyte proteome following ischemia-reperfusion injury.

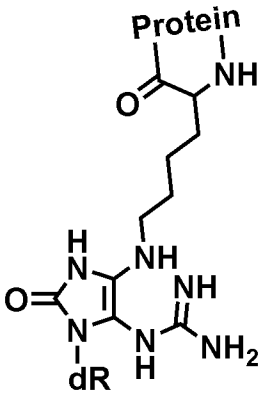
Scheme 3-1: Structures of oxidant-induced DNA-protein cross-links. Chemical structures of previously identified oxidant-induced DNA-protein cross-links including; 2-amino-3-(2-amino-10-(4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-4,11-dioxo-3,4-dihydro-4a,9a-(epiminomethanoimino)benzofuro[2,3-d]pyrimidin-6-yl)propanoic acid **([4.3.3.0] dOG-Tyr**), N6-(7-amino-1-(4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,9-dioxo-1,3,6,8-tetraazaspiro[4.4]nona-3,7-dien-4-yl)lysine (**5-Lys-Sp**), N6-(5-guanidino-1-(4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-2,3-dihydro-1H-imidazol-4-yl)lysine (**Gh-Lys**), O-(1-(4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-5-ureido-4,5-dihydro-1H-imidazole-4-carbonyl)serine (**Oxa-Ser**), N6-(7-amino-1-(4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-4,9-dioxo-1,3,6,8-tetraazaspiro[4.4]nona-2,7-dien-2-yl)lysine (**8-Lys-Sp**), N6-(2-amino-9-(4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-6-oxo-6,9-dihydro-1H-purin-8-yl)lysine (**8-Lys-dG**), and 2-amino-3-(4-hydroxy-3-((1-((2R,4R,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)methyl)phenyl)propanoic acid (**dT-Tyr**).



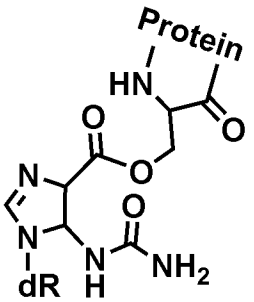
[4.3.3.0]-dOG-Tyr



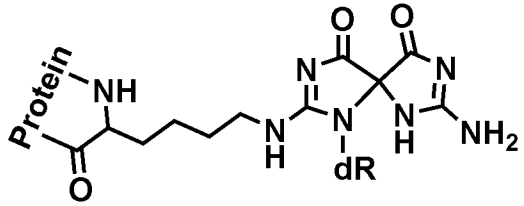
5-Lys-Sp



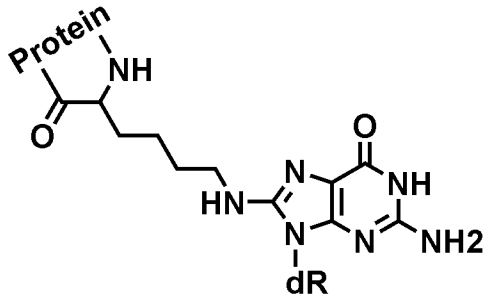
Gh-Lys



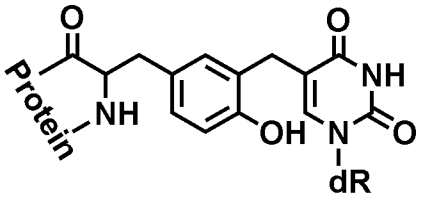
Oxa-Ser



8-Lys-Sp



8-Lys-dG



dT-Tyr

3.2. Materials and Methods

Chemicals and Reagents.

2'-Deoxythymidine, Boc-L-Tyr-OH, trifluoroacetic acid, Nuclease P_1 from *Penicillium citrinum*, and DL-dithiothreitol were purchased from Sigma-Aldrich (St. Louis, MO). Pierce™ BCA Protein Assay kit, Pierce™ Quantitative Colorimetric Peptide Assay, Pierce Quantitative Colorimetric Peptide Assay, and TMT 6-plex reagents were purchased from Thermo Fischer Scientific (Waltham, MA). Glacial acetic acid, hydrochloric acid, and acetonitrile were obtained from Fisher Scientific (Buchs, Switzerland). Mass spectrometry grade Trypsin Gold was purchased from Promega (Madison, WI). Cell lysis solution, protein precipitation solution, and RNaseA were purchased from Qiagen (Hilden, Germany). Amicon 3K filters were purchased from Millapore (Darmstadt, Germany). Omega Nanosep 10K Omega filters were obtained from PALL Life Science (Port Washington, NY). The OxiSelect™ 8-iso-prostaglandin F2α ELISA kit was purchased from Cell Biolabs (San Diego, CA). Matrin-3 and glyceraldehyde-3-phosphate dehydrogenase primary antibody and goat anti-rabbit IgG-HRP secondary antibody were purchased from Santa Cruz Biotechnology (Dallas, TX). Superoxide dismutase [Mn] mitochondrial and pyruvate dehydrogenase E1 beta subunit primary antibodies were purchased from Abcam (Cambridge, UK).

Animal safety statement

Age-matched male Sprague-Dawley rats were randomly divided into two experimental groups: left anterior descending artery (LAD) ligation surgery/reperfusion group and sham surgery control group. Animals were housed in standard rodent cages with

12-hour light-dark cycles and were given food and water *ad libitum*. All experimental procedures were approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Minnesota. All studies were conducted in accordance with the United States Department of Health and Human Services National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Animal model of myocardial infarction/reperfusion

A subset of animals within the weight range of 150–175 g underwent thoracic surgery. Anesthesia was induced with isoflurane gas (2–5% in 100% oxygen), and subsequently animals were intubated and ventilated. A thoracotomy was performed exposing the heart, and the left anterior descending (LAD) coronary artery was ligated to simulate myocardial infarction for one hour as previously described.[468] Following 1 h, the ligation was removed to restore blood flow for an additional 30 minutes to simulate reperfusion. Immediately after 30 minutes, the animals were euthanized, and the hearts excised for cardiomyocyte isolation via Langendorff reperfusion method as previously described.[469] Typical cell recovery ranged from 1.7×10^6 to 5.0×10^6 cardiomyocytes per heart, with approximately 80% cell viability. Prior to the thoracic surgery, blood was drawn from the tail vein of each animal to serve as the pre-surgery blood sample for both cardiac troponin-i (cTni) and 8-iso-prostaglandin F2α measurement assays described below. Immediately after the 30-min reperfusion, an additional blood draw occurred to serve as the post-surgery blood sample for both assays.

Troponin-I (cTni) ELISA assay

cTni levels in plasma pre- and post-LAD ligation/reperfusion was measured using the i-STAT cardiac troponin I cartridge test (Abbott, Chicago, IL) using a VetScan i-STAT 1 handheld analyzer (Abaxis, Union City, CA) following manufacturer's protocol as described previously.[470, 471]

8-iso-Prostaglandin F2α (8-iso-PGF2α) ELISA assay

8-iso-PGF2α levels in plasma pre- and post-LAD ligation/reperfusion were measured using the OxiSelect™ 8-iso-prostoglandin F2α ELISA kit following the manufacturer's protocol. In short, 50 µL plasma samples (in triplicate) were treated with 2N NaOH at 45 °C for 2 hours to hydrolyze lipoprotein or phospholipids coupled to 8-iso-PGF2α. The solutions were neutralized with 10 N HCl to pH 6.0 – 8.0. The goat anti-rabbit antibody-coated 96-well plate was incubated with 100 µL anti-8-iso-PGF2α antibody (1:1000 dilution) at 25 °C for 1 hour with gentle shaking. A serial dilution of 8-iso-PGF2α standards were prepared, and 55 µL of the standards or the prepared plasma samples were combined with 55 µL of a provided 8-iso-PGF2α-HRP conjugate. The anti-8-iso-PGF2α antibody-treated plate was incubated with the prepared standards or samples at 25 °C for 1 h with gentle shaking, followed by incubation with 100 µL HRP substrate at room temperature for 30 min with gentle shaking. The HRP reaction was quenched by adding 100 µL stop solution, and the plates were immediately analyzed on a BioTek Synergy H1 microplate reader (absorbance at 450 nm).

***5-Hydroxymethyl-2'-deoxyuracil* (5'-hm-dU)**

2'-Deoxythymidine (100 mg, 0.3875 mmol) was suspended in 25 mL 0.1M phosphate buffer (pH 7.5). Sodium persulfate (21.5 mg, 0.3875 mmol) was added, and the solution was heated at 75 °C for 4 h. The reaction mixtures were separated by semi-preparative HPLC using a Zorbex XDB C18 (250 mm x 10 mm, 5 μm) column eluted with a linear gradient of water (A) and acetonitrile (B) at 3 mL/min. Solvent composition was changed from 3% B to 5% B over 30 min, then increased to 50% over 5 min, held at 50% for 2 min, decreased to 3% over 2 minutes, and finally held at 3% for 4 min for re-equilibration. Under these conditions, 5'-hm-dU eluted at 8.7 min. ESI^+-MS/MS (5'-hm-dU): m/z = 259.2 $[M+H]^+ \rightarrow$ 143.1$[M - dR + H]^+$.

***2-amino-3-(4-hydroxy-3-((1-((2R,4R,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)methyl)phenyl)propanoic acid* (dT-Tyr)**

5'-hm-dU (180 mg, 0.4128 mmol) was dissolved in 1 mL glacial acetic acid and 6 M HCl. Boc-L-Tyr-OH (116 mg, 0.4541 mmol) was added and the solution was refluxed (110 °C) for 1 h. Following neutralization to pH 7.0 using 1 M NaOH, the precipitate was filtered and the supernatant separated by semi-preparative HPLC on a Zorbex XDB C18 (250 mm x 10 mm, 5 μm) column eluted with a linear gradient of water (A) and acetonitrile (B). The solvent composition was held at 2% B for 5 min, increased to 30% B in 5 min, then held at 30% B for 2 min, then increased to 75% B in 1 min, held at 75% B for 1 min, decreased to 2% B in 15 min, and finally held at 2% B for 10 min for re-equilibration.

Under these conditions, dT-Boc-L-Tyr-OH eluted at 19.2 min. ESI^+-MS/MS (dT-Boc-L-Tyr-OH): *m/z* = 522.3 $[M + H]^+ \rightarrow$ 406.4 $[M - dR + H]^+$, 360.4 $[M - dR - CO_2H + H]^+$. Following HPLC purification, dT-Boc-L-Tyr-OH was dissolved in 50% TFA and incubated at room temperature for 1 hour. The solution was neutralized to pH 7.0, and the supernatant was separated by the same procedure described above. Under these conditions, dT-Tyr eluted at 10 min. ESI^+-MS/MS (dT-Tyr): *m/z* = 422.1 $[M + H]^+ \rightarrow$ 306.1 $[M - dR + H]^+$ and 260.1$[M - dR - CO_2H + H]^+$.

DPC extraction from rat cardiomyocytes

Isolated rat cardiomyocytes (~ 1.0 x 10^6) were re-suspended in 1 mL of cell lysis solution (Qiagen, Hilden, Germany) and pipetted until the cells were completely lysed. The resulting lysates were treated with 40 μg RNaseA overnight at room temperature. Following incubation, each sample was treated with 750 μL of protein precipitation solution (Qiagen, Hilden, Germany) and vortexed for one minute to yield a white turbid solution. Each sample was then centrifuged at 2500 rcf for 10 minutes to pellet the protein precipitate, which was subsequently used for global proteome analyses descried below. The supernatant containing DNA and DPCs was decanted, and an equal volume of cold 100% EtOH was added to precipitate DNA/DPCs. Precipitated DNA/DPCs were washed with 70% EtOH twice and 100% EtOH, followed by re-suspending in water. DNA concentrations were determined by dG analysis as described in **Section 2.2**.

Enzymatic digestion of DPCs to tryptic peptides

DNA/DPCs (10 μg) were dissolved in 100 μL of 50 mM ammonium acetate buffer, pH 5.5 containing 5 mM $ZnCl_2$ and heated at 90 °C for 10 min to denature the DNA. Immediately after heating, the samples were chilled in an ice bath (0 °C) to prevent DNA re-annealing. Once chilled, 2 U of nuclease P_1 was added to each sample and the solutions were incubated for 4 hours at 37 ºC. After 4 hours, the pH was raised to 7.0 by adding 1 μL 1 M ammonium bicarbonate, and 120 U phosphodiesterase 1 and 22 U alkaline phosphatase (Sigma, St. Louis, MO) were added to each sample. The solutions were incubated at 37 °C overnight to complete digestion of the DNA constituent of DPCs. After digestion, the solutions were transferred to Amicon 3K filters (Millapore, Darmstadt, Germany), and the free nucleosides were removed by centrifugation at 14,000 g for 10 min.

Filter-aided sample preparation (FASP) protocol for DPC protein trypsin digestion

Protein-nucleoside conjugates were washed with 100 mM HEPES three times (14,000 g for 10 minutes) and re-suspended in 100 μL 100 mM HEPES buffer. A fresh solution of 100 mM DTT/0.1% SDS in 100 mM HEPES was prepared, and 10 μL aliquot was added to each sample, followed by heating at 55 °C for 1 hour. After heating, 5 μL of freshly prepared iodoacetamide solution (375 mM) was added, and the solution was incubated for 30 min at room temperature. Each sample was then centrifuged at 14,000 g for 10 minutes to remove excess reagents, followed by resuspension in 50 μL of 25 mM ammonium bicarbonate. Proteins were digested to tryptic peptides by incubation with 1 μg trypsin overnight at 37 °C.

FASP protocol for whole proteome trypsin digestion

The pelleted proteins from each LAD ligation/reperfusion and sham surgery sample were reconstituted in 5 mL 15 mM Tris buffer, pH 7.5/0.1% SDS and the protein concentrations were measured using the Pierce™ BCA Protein Assay kit. Equivalent amounts (100 μg) from each sample were transferred to Amicon 3K filters and washed with 100 μL 100 mM HEPES three times (14,000 g for 10 minutes). The proteins were treated with DTT and iodoacetamide as described above, followed by incubation with 1 μg trypsin overnight at 37 °C.

TMT labeling of tryptic peptides

Tryptic peptides were recovered by centrifugation at 14,000 g for 10 min and concentrated in glass MS vials to dryness. Once dry, the samples were re-suspended in 100 μL of water and peptide concentrations were determined by the Pierce Quantitative Colorimetric Peptide Assay (Thermo, Waltham, MA). For peptides from whole proteome samples, 10 μg of peptide were diluted to 35 μL of 100 mM HEPES, and the solution pH was adjusted to pH 8.0 using 100 mM NaOH. For peptides from DPC proteins, peptide amounts corresponding to 10 μg DNA were prepared as described above. To each peptide solution was added 5 μL of ACN and 10 μL of a TMT-6plex reagent (19.5 μg/μL, 40-fold excess), and the reaction mixtures were incubated at room temperature for 2 h. The reactions were quenched with 4 μL of 5% hydroxylamine for 15 minutes at room temperature.

The derivatized peptides were evaporated to dryness under reduced pressure and re-suspended in 60 µL 98/2/0.1 water/ACN/formic acid. The samples were combined and desalted by stage tip procedure as follows. Stage tips were manually packed with C18 SDX packing discs and washed with 60 µL 60/40/0.1 ACN/water/formic acid (500 rcf for 5 min) and 60 µL 98/2/0.1 water/ACN/formic acid (500 rcf for 5 minutes). The reconstituted samples were then loaded onto the tips (500 rcf for 5 minutes) and washed with 60 µL 98/2/0.1 water/ACN/formic (5000 rcf for 5 minutes) twice. Labeled peptides were eluted with 120 µL 60/40/0.1 ACN/water/ formic acid and 60 µL 80/20/0.1 ACN/water/ formic acid, then concentrated under vacuum.

nanoHPLC-NSI$^+$-MS/MS analysis of tryptic peptides

TMT-labeled peptides were reconstituted in 20 µL 0.1% formic acid and analyzed using an LTQ Orbitrap Fusion mass spectrometer (Thermo Scientific) interfaced to a Dionex Ultimate 3000 RS autosampler and an Dionex UltiMate 3000 RSLCnano system. The samples were loaded onto a pulled-tip fused silica column with a 100 µm inner diameter packed in-house with 45 cm of 5µm Luna-C18 resin (Phenomenex, Torrence, Ca) that served both as an analytical column and as a nanospray ionization emitter. HPLC flow was held at 300 nL/min. HPLC solvents were comprised of 0.1% formic acid in water (solvent A) and 0.1% formic acid in acetonitrile (solvent B). Solvent composition was held at 5% solvent B for 17 min, followed by an increase to 7% over 2 min, 25% over 100 min, 60% over 20 min, and finally to 95% over 1 min. The solvent composition was held at 95% for 8 min, and returned to 5% over 1 minute and re-equilibration for 28 min.

The mass spectrometer was operated in a data dependent mode with dynamic exclusion enabled (repeat count: 1, exclusion duration: 15 sec). Every scan cycle, one full MS scan (*m/z* 320 to 2000) was collected at a resolution of 120,000 at an AGC target value of $2x10^5$, followed by MS^2 scans of as many dependent scans as possible within a cycle time of 3 seconds with HCD (normalized collision energy = 38%, isolation width = 0.7 *m/z*, resolution = 30,000) at an AGC target value of $1x10^5$. Ions with a charge state of +1 were excluded from the analysis.

Protein identification

Peptide sequences were determined from their MS^2 spectra using Proteome Discover 2.1 software (Thermo, Waltham, MA) and a SwissProt rat proteome database.[472] Precursor ion *m/z* mass tolerance was set at ± 10 ppm, fragment ion *m/z* tolerance at ± 0.06 Da, and up to two missed cleavages were allowed. Static modifications included carbamidomethylation (+57.02146 Da) of cysteine and TMT 6-plex (+219.163 Da) groups on lysines, tyrosines and protein N-termini. Dynamic modifications included oxidation (+15.99492 Da, M) and deamidation (+0.98402 Da, N). Identified peptides were filtered using a 1% false discovery rate and at least 1 unique peptide fragment. Isotope distributions for all six TMT reporter regions, provided by the supplier, were included in the analysis.

Relative protein quantitation

For relative quantification of DPC proteins and free proteins in cardiomyocytes from LAD surgery and sham surgery groups, reporter ion intensities (*m/z* 126.1277, 127.1247, 128.1344, 129.1314, 130.1411, 131.1381) were extracted using Proteome

Discoverer 2.1 software. All quantitative values were rejected if not all quantitative channels were present, and we did not replace missing quantitative values with the minimum intensity. The mass tolerance for reporter ion extraction was 0.01 Da, and the signal to noise ratio of the reporter ion was at least 10.

Bioinformatic analysis of quantitative proteomics data

Whole proteome and DPC protein abundance data were processed by the Ingenuity Pathway Analysis (IPA, Qiagen, Hilden, Germany) and Cluster 3.0. To develop protein interaction networks, protein TMT abundance data were processed using a core expression analysis measuring expression fold change. Whole proteome and DPC TMT abundance data was further processed using Ingenuity Pathway Analysis (IPA) comparison analysis to identify shared protein interaction networks (QIAGEN Inc., https://www.qiagenbioinformatics.com/ products/ingenuity-pathway-analysis). Clustering analysis of the protein TMT abundance data was analyzed using a hierarchical clustering algorithm (uncentered correlation).[473]

Quantitation of dT-Tyr conjugates

DPC-containing DNA (10 μg) isolated from rat cardiomyocytes as described above was dissolved in 100 μL of 10 mM Tris, pH 7.5 containing 8 Units proteinase K (10 μL, New England Biolabs, Ipswich, MA) and 50 μg of pronase (Sigma, St. Louis, MO) at 37 °C overnight to digest the protein constituents of DPCs to single amino acids. DNA was enzymatically digested to nucleosides as follows. The solutions were concentrated to dryness, and samples were re-suspended in 100 μL 10 mM Tris, pH 5.5 with 5 mM $ZnCl_2$.

Each sample was supplemented with 2.5 U of $NucP_1$ and incubated at 37 °C overnight. Next morning, the solution pH was raised to pH 7.0 by adding 1 µL 1 M ammonium bicarbonate, followed by adding 15 µL of 100 mM $MgCl_2$, 35 U DNase1, and 160 U of PDE1 (total volume 150 µL) and incubated for 4 h at 37 °C. Each sample was then spiked with 22 U of alkaline phosphatase and incubated for an additional at 37 °C overnight.

The digests were spiked with dT[$^{15}N_1$,$^{13}C_9$]-Tyr (internal standard for mass spectrometry, 200 fmol), followed by ultrafiltration through Nanosep 10K filters at 5000 g for 10 minutes. The filters were further washed with 100 µL 0.1% formic acid in water and 100 µL 0.1% formic acid in ACN. The filtrates were concentrated to dryness, re-suspended in 100 µL 0.1% acetic acid, and enriched by offline HPLC purification as follows. An Agilent Technologies HPLC system (1100 model) was used incorporating a diode array detector, an autosampler, and a fraction collector. HPLC solvents were comprised of 0.1% acetic acid in water (solvent A) and 0.1% acetic acid in acetonitrile (solvent B). The samples were loaded on a Supelcosil LC-18 (4.6 × 250 mm, 5 µm) column (Thermo Scientific, Waltham, MA) and eluted at a flow rate of 1 mL/min. Solvent composition was held at 2% B over 5 min, increased to 30% over 7 min, further to 75% over 1 min, held at 75% for 2 min, decreased to 2% over 1 min, and finally re-equilibrated at 2% B for 14 min. Fractions containing dT-Tyr and its isotopically labeled internal standard (12.9 – 13.9 min) were collected, dried under vacuum, and reconstituted in 0.1% formic acid (16 µL) for quantitative analysis. Due to interfering peaks, additional HPLC purification of the collected analyte and internal standard using the same method was required for some samples. To exclude the possibility of analyte carryover between runs, a blank sample (100

μL of 0.1% acetic acid in water) spiked with 200 fmol internal standard was analyzed after every third sample.

Quantitative analysis of dT-Tyr was conducted using a Dionex UltiMate 3000 RSLC nanoHPLC system (Thermo Scientific, Waltham, MA) interfaced to a TSQ Quantiva mass spectrometer (Thermo Scientific, Waltham, MA). The samples were loaded onto a pulled-tip fused silica column with a 100 μm inner diameter packed in-house with 15 cm of Synergi 4 μm Hydro RP 80A resin (Phenomenex, Torrence, Ca) that served both as a resolving column and as a nanospray ionization emitter. HPLC solvents were comprised of 0.1% formic acid in water (solvent A) and 0.1% formic acid in acetonitrile (solvent B). dT-Tyr was eluted with a gradient of 1% solvent B over 6 min at a flow rate of 1.0 μL/min, followed by an increase to 30% over 10 min at a flow rate of 0.3 μL/min, then held at 30% over 2 min, then increased to 75% over 2 min, decreased to 1% over 1 min, then finally re-equilibrated at 1% for 11 min at 1 μL/min. Under these conditions, dT-Tyr and its internal standard (dT-[$^{15}N_1$, $^{13}C_9$]Tyr) eluted at ~ 13.8 min. Electrospray ionization was achieved at a spray voltage of 2800 V and a capillary temperature of 350 °C. Collision induced dissociation was performed with Ar as a collision gas (1.0 mTorr) at a collision energy of 23 V. MS/MS parameters were optimized for maximum response during infusion of a standard solution of dT-[$^{15}N_1$, $^{13}C_9$]Tyr.

nanoLC-ESI$^+$-MS/MS analysis of dT-Tyr was performed in the selected reaction monitoring mode by monitoring the neutral loss of dR and the loss of dR and CO_2 from protonated molecules of the analyte (*m/z* 422.2 $[M + H]^+ \rightarrow$ 306.1 $[M + H - dR]^+$, *m/z* 422.2 $[M + H]^+ \rightarrow$ 260.1) and the corresponding mass transitions corresponding to $^{15}N_1$,$^{13}C_9$- labeled internal standard (*m/z* 432.1 $[M + H]^+ \rightarrow$ 316.1 $[M + H - dR]^+$, *m/z* 432.2

$[M + H]^+ \rightarrow 269.1$). Analyte concentrations were determined using the relative response ratios calculated from HPLC-ESI$^+$-MS/MS peak areas in extracted ion chromatograms corresponding to dT-Tyr and its internal standard.

nanoLC-ESI$^+$-MS/MS method validation

Calf thymus DNA (CT DNA, 50 μg, in triplicate) was spiked with increasing amounts of dT-Tyr analyte (2-500 fmol) and 200 fmol dT-[$^{15}N_1$,$^{13}C_9$]Tyr. Each sample was processed as described above. The linearity of each validation sample was determined by plotting the observed peak area ratio (analyte/internal standard) to the expected peak area ratio. The limit of detection was defined as the analyte signal at least 3 times the response as compared to the background. Accuracy and precision of the method was calculated for three samples (200 fmol, high point; 50 fmol, middle point; 10 fmol, low point). Accuracy was calculated by subtracting the average observed peak area ratio (analyte/internal standard) by the expected peak area ratio, followed by dividing the difference by the expected peak area ratio multiplied by 100. Accuracy was expressed as the % difference. Precision was calculated by dividing the standard deviation between samples by the mean observed peak area multiplied by 100, expressed as % coefficient of variation (%CV).

Western Blot Analysis

Cardiomyocyte proteins (50 μg) were resolved by NuPAGE Novex 12% Bis-Tris gels (Invitrogen, Carlsbad, CA) and transferred to Invitrolon PVDF filter paper membranes (0.45 μm pore size, Life Technologies, Carlsbad, CA). The membranes were blocked for 1 h in Tris-buffered saline-Tween 20 (TBST) containing 5% (w/v) blotting grade blocker

(BioRad, Hercules, CA). Following blocking, the membranes were incubated with the primary antibodies against matrin-3, superoxide dismutase [Mn], mitochondrial (SOD2), pyruvate dehydrogenase E1 component subunit beta, mitochondrial (PDH-1Eβ), troponin-T (cTni), and actin overnight at 4 °C. The membranes were rinsed with TBST and incubated with the corresponding alkaline phosphatase-conjugated antibody for 1 h at room temperature. The membranes were developed using the Clarity Western ECL substrate (BioRad) and visualized on an Odyssey 2800 Fc imager (Li-COR, Lincoln, NE) with Image Studio version 3.1 software (Li-COR). To confirm the identity of the protein constituents from cardiomyocyte DPCs, 5 μg of DNA from one LAD ligation/reperfusion and sham surgery samples were enzymatically digested to yield a nucleoside-protein conjugate. Nucleoside-protein conjugates were resolved and immunoblotted as described above.

3.3. Results

3.3.1. Ischemia-reperfusion injury increases plasma troponin I levels and induces ROS

To confirm that the LAD ligation procedure successfully induced myocardial infarction, plasma levels of cardiac troponin I (cTnI) were measured using a sensitive ELISA test.[474, 475] cTnI is a useful biomarker of MI commonly used in the clinic to diagnose MI and acute coronary syndrome.[476, 477] cTnI levels in all plasma samples collected before the thoracotomy were below the limit of detection of the assay (**Table 3-1**). The animals that underwent a sham surgery (thoracotomy only) experienced a modest increase in plasma cTnI levels to 4.4-6.6 ng/mL respectively (**Table 3-1**). In contrast, plasma levels of

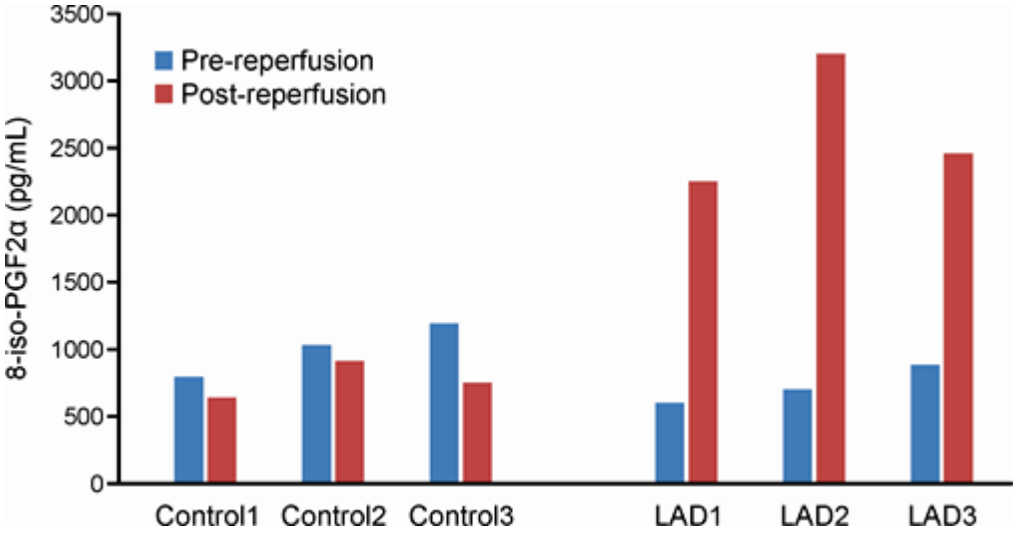
cTnI in animals that underwent LAD ligation/reperfusion procedure were above 50 ng/mL, confirming that they experienced an MI (**Table 3-1**).

To determine whether the ischemia/reperfusion procedure led to an increase in ROS, plasma levels of 8-iso-PGF2α were measured pre- and post-procedure using the OxiSelect™ 8-iso-prostoglandin F2α ELISA assay.[478-480] 8-iso-PGF2α is an isoprostane produced by ROS-mediated peroxidation of arachidonic acid in membrane phospholipids and is commonly used as a biomarker of oxidative stress *in vivo.*[478-480] It is important to note that prior to the surgeries, animals where anesthetization under 100% oxygen, potentially exposing them to hyperoxic conditions and an artificial increase in ROS levels. Therefore, the baseline levels of 8-iso-PGF2α could have been elevated prior to the surgery. We found that 8-iso-PGF2α levels slightly decreased post-thoracotomy in control group that underwent sham surgery (Controls 1 – 3) (**Figure 3-1**). In contrast, animals that underwent LAD ligation/reperfusion procedure (LAD 1 – 3) experienced a 273.8%, 357.1%, and 178.3% increase in 8-iso-PGF2α levels post-surgery compared to the pre-procedure levels, indicative of increased oxidant levels in this group (**Figure 3-1**).

Table 3-1: Plasma cTni measurements pre- and post-surgery to confirm myocardial infarction. The cTni levels after LAD ligation/reperfusion was above the limit of detection of the assay (50 ng/mL), while the average cTni levels after the sham surgery was 5.82 ± 1.03 ng/mL.

Experiment	Pre-Surgery cTni (ng/mL)	Post-Surgery cTni (ng/mL)
LAD 1	ND	>50
LAD 2	ND	>50
LAD 3	ND	>50
Control 1	ND	6.57
Control 2	ND	6.52
Control 3	ND	4.37

Figure 3-1: Plasma 8-iso-PGF2α measurements. Plasma 8-iso-PGF2α levels were measured pre- and post-surgery using the OxiSelect™ 8-iso-prostoglandin F2α ELISA kit. A standard curve of 200,000 ng/mL, 50,000 ng/mL, 12,500 ng/mL, 3125 ng/mL, 781ng/mL, 195 ng/mL, and 49 ng/mL was analyzed by a "log inhibitor vs response, variable slope with 4-parameters" equation to determine 8-iso-PGF2α concentration. The animals that underwent LAD ligation/reperfusion (LAD 1, 2, and 3) had a significant increase (p value < 0.05) in 8-iso-PGF2α post-surgery as compared to pre-surgery. No significant increase in 8-iso-PGF2α levels was detected in the animals that underwent the sham surgery (Control 1, 2, and 3).



3.3.2. Total DPC numbers are increased following ischemia-reperfusion injury

To allow for absolute quantification of ROS-induced DPCs in rat cardiomyocytes, an isotope dilution tandem mass spectrometry assay was developed for dT-Tyr conjugates (**Scheme 3-2**). This type of ROS-induced cross-linking involves radical-mediated reactions between the 5'-methyl of thymidine in DNA and the 3-position of tyrosine residue in proteins (**Scheme 3-2**).[362, 363, 481, 482] DPCs extracted from rat cardiomyocytes were enzymatically digested with nucleases and proteases to afford dT-Tyr and analyzed by isotope dilution HPLC-ESI$^+$-MS/MS allowing for accurate and precise quantification of hydroxyl radical-induced DNA-protein conjugates.

Authentic standards of dT-Tyr and dT-[$^{15}N_1$,$^{13}C_9$]Tyr were prepared in our laboratory using published methodologies.[483] These standards were used to develop a sensitive and precise nanoLC-ESI$^+$-MS/MS method for dT-Tyr using a triple quadrupole mass spectrometer. In our approach, DPC-containing DNA is spiked with known amounts of internal standard and enzymatically digested to free nucleosides. The analyte and its internal standard are enriched by offline HPLC, followed by quantitative analysis using HPLC-ESI$^+$-MS/MS. Selected reaction monitoring transitions included the neutral loss of dR and the simultaneous loss of dR and CO_2 from protonated molecules of the analyte (*m/z* 422.2 $[M + H]^+ \rightarrow$ 306.1 $[M + H - dR]^+$, *m/z* 422.2 $[M + H]^+ \rightarrow$ 260.1) and the $^{15}N_1$,$^{13}C_9$-isotopically labelled internal standard *(m/z* 432.1 $[M + H]^+ \rightarrow$ 316.1 $[M + H - dR]^+$, *m/z* 432.2 $[M + H]^+ \rightarrow$ 269.1) (**Figure 3-2A**). The new HPLC-ESI$^+$-MS/MS method was validated by analyzing blank DNA samples spiked with known amounts of dT-Tyr analyte and internal standard. HPLC-ESI$^+$-MS/MS limit of detection (LOD) was calculated as 250 amol of dT-Tyr on column, while the limit of quantification was 2 fmol on column (**Figure**

3-3). The accuracy (% difference) of the method was calculated as 3.0%, 4.0%, and 20.0% for the high, middle, and low validation points. The precision of the method was calculated as 4.1, 4.2, and 0.0 % CV for the high, middle, and low validation points respectively.

The new method was used to quantify radical-induced DPCs after LAD ligation/ reperfusion compared to sham surgery (3 per group). We found that dT-Tyr amounts were significantly increased in cardiomyocytes of rats that underwent LAD ligation/reperfusion as compared to the sham group (35.7 ± 2.2 vs 18.2 ± 1.5 dT-Tyr adducts per 10^8 nucleotides, respectively, p-value < 0.01) (**Figure 3-2B**). These findings support our hypothesis that ischemia-reperfusion injury increases the levels of hydroxyl radical-mediated DPCs in cardiomyocyte DNA.

Scheme 3-2: Proposed mechanism of hydroxyl radical-induced DPC formation between thymidine in DNA and tyrosine residue in proteins. Hydroxyl radicals abstract a hydrogen from the 5-methyl position of thymidine to yield a reactive thymidine radical (**19**), which undergoes a one-electron addition to the 3-position of a tyrosine to yield a stable methylene linkage. Subsequent hydrogen abstraction from the 3-position of tyrosine (**20**) re-aromatizes the phenol ring to yield the stable DNA-protein cross-link (**21**).

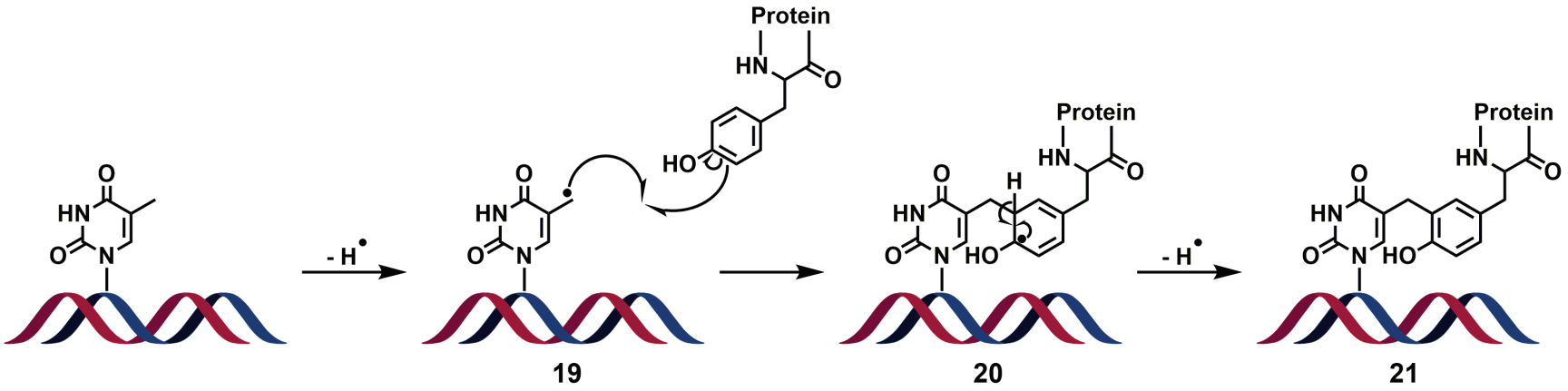


Figure 3-2: Representative nanoHPLC-ESI^{+}-MS/MS trace of dT-Tyr observed in rat cardiomyocytes following LAD ligation/reperfusion surgery (**A**) and quantitative data for dT-Tyr adducts formed in rat cardiomyocytes after LAD ligation/reperfusion or sham surgery (**B**).

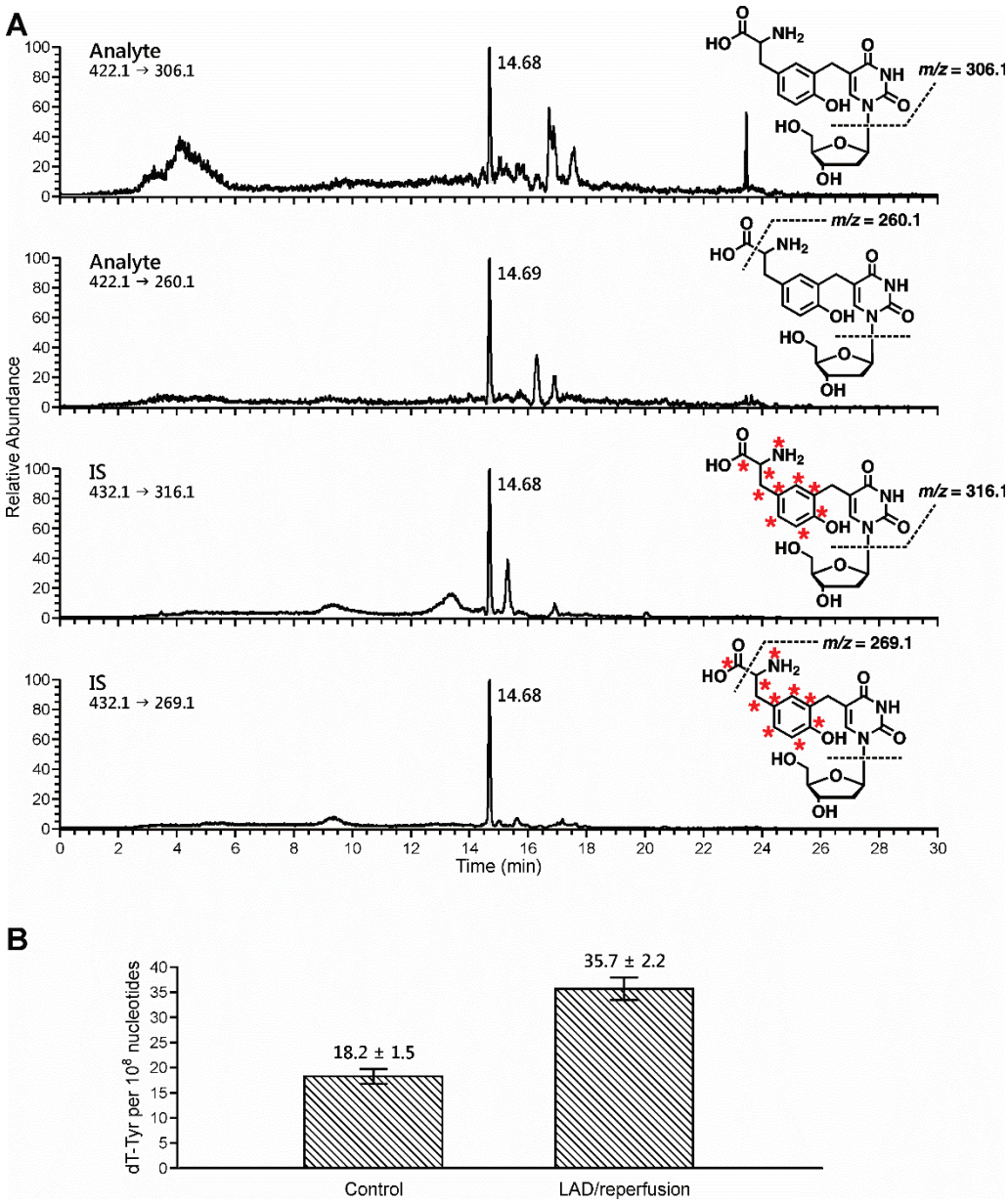
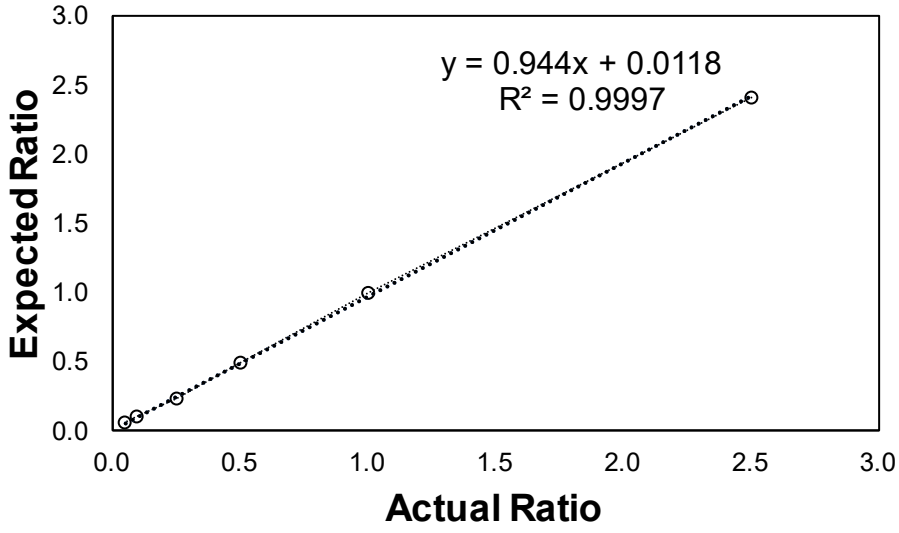


Figure 3-3: Validation of the developed dT-Tyr nanoLC-ESI^{+}-MS/MS assay. Calf thymus DNA (50 µg, in triplicate) was spiked with dT-Tyr analyte (500, 200, 100, 50, 20, 10, or 2 fmol) and 200 fmol dT-Tyr ($^{15}N_1$,$^{13}C_9$) IS. Each sample was processed for dT-Tyr enrichment as described in the Materials and Methods (**Section 3.2**). The linearity of each validation replicate was determined by plotting the observed peak area ratio (analyte/internal standard) to the expected peak area ratio. The limit of detection was defined as the analyte signal at least 3 times the response as compared to the background.



3.3.3. Quantitative proteomic analysis of the protein constituents of cardiomyocyte DPCs

We next aimed to identify and quantify the proteins participating in DPC formation following LAD ligation/reperfusion surgery. Following DPC isolation from control and LAD ligation/reperfusion rat cardiomyocytes, protein constituents were analyzed by TMT quantitative proteomics technique.[484] In brief, each of the six samples (3 LAD ligation/reperfusion, 3 sham surgery) was labeled with a unique isotope tag. The samples were combined and analyzed by nanoLC-ESI-MS/MS on an Orbitrap Fusion mass spectrometer. MS/MS sequencing of tryptic peptides was used to determine their amino acid sequences, and the proteins were identified using Proteome Discoverer 2.1 software.[485] Relative quantitation of the proteins was achieved using ion intensities of reporter ions (*m/z* 126.1277, 127.1247, 128.1344, 129.1314, 130.1411, 131.1381) originating from their tryptic peptides. The relative abundances of proteins in LAD ligation/reperfusion animals as compared to the sham group was determined by comparing the grouped reporter ion intensities of 126.1277, 128.1344, and 131.1381 (LAD 1-3) to the grouped reporter ion intensities of corresponding to control group (*m/z* 127.1247, 129.1314, and 130.1411 (controls 1-3).

The relative abundances of proteins covalently bound to DNA following LAD ligation/reperfusion were determined by comparing the grouped reporter ion intensities of 128.1344, 129.1314 and 131.1381 (LAD 1-3) to the grouped reporter ion intensities of 126.1277, 127.1247, and 130.1411 (control animals 1-3). Using these criteria, a total of 90 proteins were identified in all six samples (**Table 3-2**). The majority of the identified DPC protein constituents (80, 88.9%) increased in abundance at least 1.2-fold after LAD

ligation/reperfusion as compared to sham surgery, with 21 (23.3%) of these proteins increasing in abundance at least 5-fold (**Table 3-2**). Proteins that showed the greatest change included myosins, heat shock proteins, Cu-Zn superoxide dismutases, thioredoxin, actin, and mitochondrial creatinine kinase. Of the remaining DPC proteins, 8 proteins (8.9%) did not change in abundance and one protein (1.1%, lysosome-associated membrane glycoprotein 1) decreased in abundance (**Table 3-2**).

Table 3-2: Identification and quantification of proteins irreversible trapped on DNA in cardiomyocytes of rats subjected to ischemia-reperfusion or sham surgery.

Accession	Description of Protein	% Coverage	# Peptides	# PSMs	# Unique Peptides	MW [kDa]	Fold Change (LAD/Sham)
P13437	3-ketoacyl-CoA thiolase, mitochondrial	5.54	2	2	2	41.84	12.38
P49432	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	13.65	5	10	5	38.96	9.28
Q64591	2,4-dienoyl-CoA reductase, mitochondrial	8.06	3	4	3	36.11	8.95
P29117	Peptidyl-prolyl cis-trans isomerase F, mitochondrial	6.80	2	2	2	21.80	8.70
P11980	Pyruvate kinase PKM	3.58	2	2	2	57.78	8.65
Q9Z0V6	Thioredoxin-dependent peroxide reductase, mitochondrial	21.40	5	7	5	28.28	7.77
P42123	L-lactate dehydrogenase B chain	23.95	8	11	7	36.59	7.38
P35434	ATP synthase subunit delta, mitochondrial	13.69	2	11	2	17.58	6.98
P13086	Succinyl-CoA ligase [ADP/GDP-forming] subunit alpha, mitochondrial	9.54	3	4	3	36.13	6.66
P04642	L-lactate dehydrogenase A chain	9.64	3	5	2	36.43	6.46
P14604	Enoyl-CoA hydratase, mitochondrial	32.41	9	22	9	31.50	6.42

P26284	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	5.90	3	3	3	43.20	6.20
P07895	Superoxide dismutase [Mn], mitochondrial	13.06	3	5	3	24.66	5.95
Q8VHF5	Citrate synthase, mitochondrial	17.38	8	13	8	51.83	5.95
P62898	Cytochrome c, somatic	43.81	5	8	5	11.60	5.94
P02563	Myosin-6	48.19	118	457	39	223.37	5.90
P02564	Myosin-7	42.38	96	355	17	222.95	5.87
P15650	Long-chain specific acyl-CoA dehydrogenase, mitochondrial	39.07	17	60	17	47.84	5.60
O35854	Branched-chain-amino-acid aminotransferase, mitochondrial	4.58	2	2	2	44.25	5.49
P18418	Calreticulin	4.81	2	2	2	47.97	5.47
O88767	Protein deglycase DJ-1	8.47	2	2	2	19.96	5.11
P51868	Calsequestrin-2	10.17	4	5	4	47.81	5.07
P04636	Malate dehydrogenase, mitochondrial	51.18	16	59	16	35.66	4.91
P00507	Aspartate aminotransferase, mitochondrial	39.07	22	54	22	47.28	4.72
Q01205	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial	7.27	4	4	4	48.89	4.69
P19804	Nucleoside diphosphate kinase B	28.95	4	5	4	17.27	4.55
P07483	Fatty acid-binding protein, heart	57.89	8	34	8	14.77	4.53

P31044	Phosphatidylethanolamine-binding protein 1	25.13	3	8	3	20.79	4.45
P04785	Protein disulfide-isomerase	8.64	4	4	4	56.92	4.40
P26772	10 kDa heat shock protein, mitochondrial	48.04	5	9	5	10.90	4.16
P08461	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial	14.87	9	25	9	67.12	4.07
P16409	Myosin light chain 3	73.50	15	51	12	22.14	3.99
P70623	Fatty acid-binding protein, adipocyte	21.21	3	3	3	14.70	3.98
P00564	Creatine kinase M-type	35.43	14	32	12	43.02	3.88
Q9JM53	Apoptosis-inducing factor 1, mitochondrial	7.84	4	5	4	66.68	3.82
Q4FZU2	Keratin, type II cytoskeletal 6A	4.35	3	6	2	59.21	3.77
P48500	Triosephosphate isomerase	20.08	5	7	5	26.83	3.75
P08503	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	6.41	3	4	3	46.53	3.71
O88989	Malate dehydrogenase, cytoplasmic	29.34	11	19	11	36.46	3.70
Q62812	Myosin-9	1.94	4	4	3	226.20	3.69
P04692	Tropomyosin alpha-1 chain	88.73	51	326	32	32.66	3.59
P13221	Aspartate aminotransferase, cytoplasmic	26.15	9	12	9	46.40	3.58
P09605	Creatine kinase S-type, mitochondrial	33.41	16	124	15	47.36	3.46

P56741	Myosin-binding protein C, cardiac-type	17.43	20	45	20	140.67	3.46
O35796	Complement component 1 Q subcomponent-binding protein, mitochondrial	11.83	2	2	2	30.98	3.45
P58775	Tropomyosin beta chain	48.24	23	125	4	32.82	3.43
P05065	Fructose-bisphosphate aldolase A	10.16	4	5	4	39.33	3.38
P15999	ATP synthase subunit alpha, mitochondrial	26.94	13	29	13	59.72	3.30
P06761	78 kDa glucose-regulated protein	8.72	5	7	5	72.30	3.29
P23965	Enoyl-CoA delta isomerase 1, mitochondrial	14.53	4	5	4	32.23	3.26
P11507	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	5.27	5	16	5	114.69	3.24
P56571	ES1 protein homolog, mitochondrial	19.17	6	10	6	28.16	3.18
Q68FU3	Electron transfer flavoprotein subunit beta	12.94	3	5	3	27.67	3.18
P08733	Myosin regulatory light chain 2, ventricular/cardiac muscle isoform	59.64	10	40	10	18.87	3.15
P02770	Serum albumin	3.45	2	7	2	68.69	3.10
P48721	Stress-70 protein, mitochondrial	7.36	4	4	3	73.81	2.97
Q62651	Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial	25.99	9	11	9	36.15	2.91
P07632	Superoxide dismutase [Cu-Zn]	22.73	3	5	3	15.90	2.77
P13803	Electron transfer flavoprotein subunit alpha, mitochondrial	15.02	4	6	4	34.93	2.75

Q5XIF3	NADH dehydrogenase [ubiquinone] iron-sulfur protein 4, mitochondrial	31.43	5	9	5	19.73	2.71
Q6P6R2	Dihydrolipoyl dehydrogenase, mitochondrial	20.04	9	15	9	54.00	2.69
P18886	Carnitine O-palmitoyltransferase 2, mitochondrial	3.50	3	3	3	74.06	2.65
D3ZHA0	Filamin-C	0.77	2	2	2	290.80	2.52
P63018	Heat shock cognate 71 kDa protein	9.44	7	11	6	70.83	2.51
Q9ER34	Aconitate hydratase, mitochondrial	19.87	15	33	15	85.38	2.38
P16290	Phosphoglycerate mutase 2	17.00	5	5	5	28.74	2.36
P28042	Single-stranded DNA-binding protein, mitochondrial	11.26	2	2	2	17.44	2.21
Q6IFW6	Keratin, type I cytoskeletal 10	9.70	5	10	4	56.47	2.17
Q6IMF3	Keratin, type II cytoskeletal 1	6.24	4	6	3	64.79	2.08
P11240	Cytochrome c oxidase subunit 5A, mitochondrial	21.23	3	3	3	16.12	2.03
Q6IFV3	Keratin, type I cytoskeletal 15	3.58	3	3	2	48.84	1.95
Q9ESS6	Basal cell adhesion molecule	7.53	4	5	4	67.47	1.92
P62738	Actin, aortic smooth muscle	47.21	16	65	6	41.98	1.78
P01946	Hemoglobin subunit alpha-1/2	15.49	2	3	2	15.32	1.75
P04762	Catalase	5.88	3	5	3	59.72	1.65
P56574	Isocitrate dehydrogenase [NADP], mitochondrial	5.75	2	3	2	50.94	1.59
P10719	ATP synthase subunit beta, mitochondrial	17.77	7	8	7	56.32	1.56

P32551	Cytochrome b-c1 complex subunit 2, mitochondrial	9.07	3	3	3	48.37	1.42
Q920L2	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial	2.90	2	3	2	71.57	1.42
Q6UPE1	Electron transfer flavoprotein-ubiquinone oxidoreductase, mitochondrial	3.73	2	2	2	68.16	1.41
Q05962	ADP/ATP translocase 1	16.11	5	6	3	32.97	1.38
Q64119	Myosin light polypeptide 6	21.85	4	8	1	16.96	1.14
P50339	Chymase	9.72	2	2	2	27.55	1.11
P51740	Intestinal-type alkaline phosphatase 2	7.08	4	14	2	59.76	1.09
P15693	Intestinal-type alkaline phosphatase 1	8.33	4	21	2	58.37	1.07
P26453	Basigin	7.99	3	5	3	42.41	1.02
P02091	Hemoglobin subunit beta-1	27.21	3	4	3	15.97	1.00
P21961	Mast cell carboxypeptidase A (Fragment)	12.14	6	7	6	47.91	0.91
Q01129	Decorin	7.34	3	4	3	39.78	0.91
P14562	Lysosome-associated membrane glycoprotein 1	4.91	2	3	2	43.94	0.83

The proteomics data for DNA-conjugated proteins was further analyzed using the Cluster 3.0 software and IPA. A subset of 25 identified DPC proteins, consisting of the most abundant proteins, underwent hierarchical clustering analysis. Significant differences in protein abundances were observed between the LAD ligation/reperfusion replicates as compared to the sham surgery group (**Figure 3-4**). Hierarchical clustering of the 25 proteins revealed tight grouping between the LAD ligation/reperfusion replicates and the sham surgery replicates, but not between the two conditions (**Figure 3-4**). IPA core analysis of the DPC proteins developed 7 protein interaction networks; organismal injury and abnormality (score 59, 26 proteins), cardiovascular system development and function (score 33, 17 proteins), free radical scavenging (score 28, 15 proteins), developmental disorder/hereditary disorder (score 26, 14 proteins), energy production or lipid metabolism (score 24, 13 proteins), and cell morphology and cellular function and maintenance (score 21, 12 proteins). Relative protein quantitation results for SOD2, PDH-1Eβ, and actin were further confirmed by western blot analysis using commercially available antibodies (**Figure 3-5**). Overall, our results indicate that ischemia/reperfusion injury drastically increases the concentrations of proteins covalently attached to DNA, potentially leading to cell death and tissue damage.

Figure 3-4: Heat map analysis of the TMT abundance data for 23 representative proteins identified from the DPC samples. Proteins analyzed include sarcomere proteins (8), ROS scavenging proteins (8), and the remaining most abundant proteins (7).

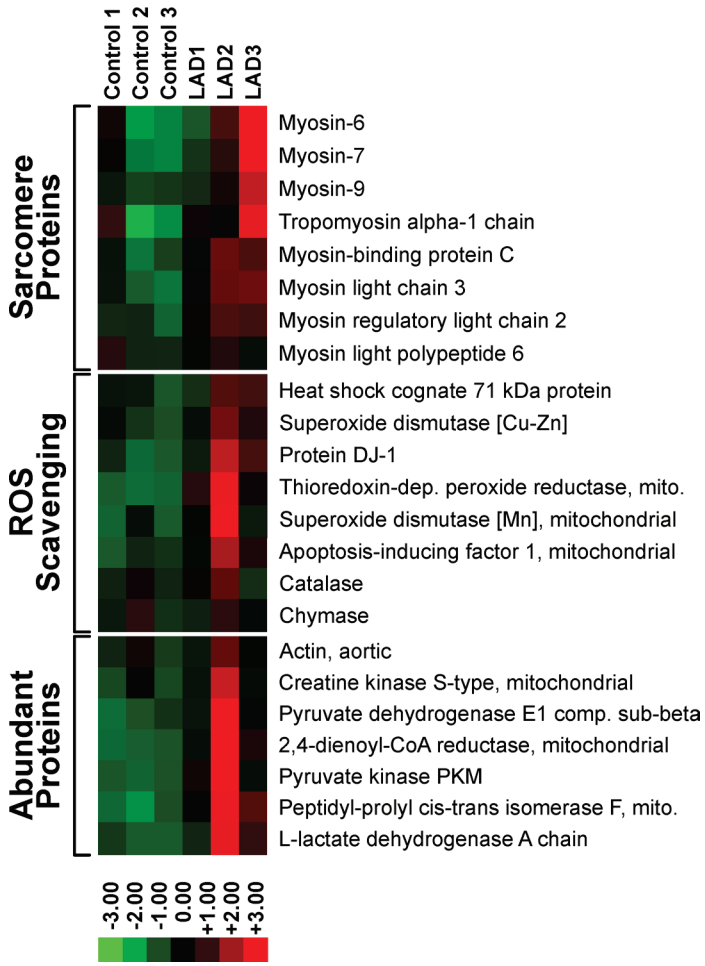
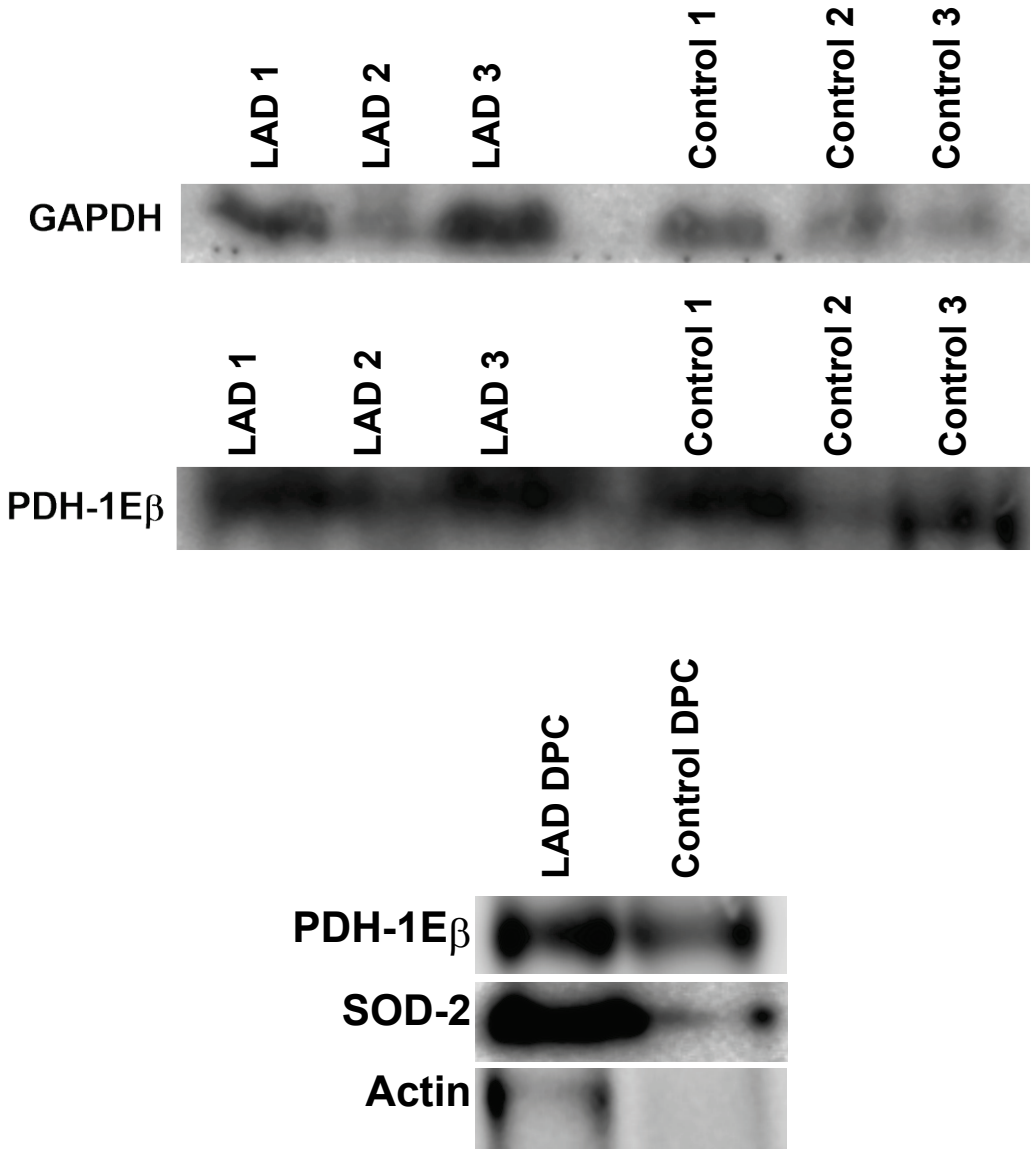


Figure 3-5: Western blot confirmation of proteins identified in the whole-proteome analysis and DPC analysis after LAD ligation/reperfusion compared to sham. For whole proteome analysis, 50 μg of protein from each sample was immunoblotted against GAPDH, PDH-1Eβ, troponin-T, and actninin primay antibody. For DPC analysis, 5 μg of DNA from a LAD ligation/reperfusion and sham surgery sample was enzymatically digested to yield protein-nucleoside conjugates that were immunoblotted against SOD2 and PDH-1Eβ.



3.3.4. Quantitative proteomic analysis of cardiomyocyte proteome

To determine whether the observed increases in DNA-protein cross-linking are linked to global changes in protein abundance in response to LAD ligation/reperfusion injury, the entire cardiomyocyte proteome from control and MI/reperfusion groups were analyzed using the same TMT quantitative proteomics technique described above [484]. A total of 359 proteins were identified in all six samples (**Table 3-3**). Of these, the majority (N = 288, 80.2%) did not change in abundance after LAD ligation/reperfusion as compared to the sham surgery (**Table 3-3**). Of the remaining 71 proteins, 48 proteins were found to increase in abundance at least 1.2-fold after LAD ligation/reperfusion and 23 proteins decreased in abundance under the same conditions (**Table 3-3**).

The majority of the proteins that increased in abundance following LAD ligation/ reperfusion (N =20, 41.7%) were mitochondrial proteins including malate dehydrogenase, enoyl-CoA hydratase, cytochrome C oxidase, and NADH dehydrogenase. This is not unexpected, as the sudden switch from hypoxic conditions to normoxic conditions when the blood flow to the affected area is restored increases the expression of proteins required for aerobic respiration [367]. The remaining proteins were categorized as membrane/extracellular proteins (14, 29.2%), sarcomere proteins (5, 10.4%), cytoplasmic proteins (4, 8.3%), nuclear proteins (4, 8.3%), and endoplasmic reticulum proteins (1, 2.1%). Among the proteins that decreased in abundance following LAD ligation/reperfusion, 39.1% (9) were nuclear, followed by cytoplasmic (6, 26.1%), mitochondrial (3, 13.0%), cell membrane (3, 13.0%), and endoplasmic reticulum (2, 8.7%). Relative protein quantitation results for GAPDH and PDH-1Eβ were further confirmed by western blot analysis using commercially available antibodies (**Figure 3-5**).

Table 3-3: Table of identified proteins from the cardiomyocyte proteome of rats subjected to ischemia-reperfusion or sham surgery.

Accession	Description	Coverage	# Peptides	# PSMs	# Unique Peptides	MW [kDa]	Fold Change (LAD/Sham)
Q6IMF3	Keratin, type II cytoskeletal 1	5.12	3	5	3	64.79	2.90
Q5M7U6	Actin-related protein 2	1.78	2	5	1	44.71	2.17
P50753	Troponin T, cardiac muscle	20.40	8	11	8	35.71	1.68
P04636	Malate dehydrogenase, mitochondrial	60.95	18	91	18	35.66	1.63
P14604	Enoyl-CoA hydratase, mitochondrial	25.17	8	14	8	31.50	1.61
P51638	Caveolin-3	11.92	3	4	1	17.39	1.58
P12847	Myosin-3	12.78	33	227	1	223.72	1.53
P02091	Hemoglobin subunit beta-1	77.55	11	92	3	15.97	1.53
P23693	Troponin I, cardiac muscle	15.64	5	12	5	24.15	1.49
Q6PCU8	NADH dehydrogenase [ubiquinone] flavoprotein 3, mitochondrial	29.63	3	7	3	11.93	1.48
P01946	Hemoglobin subunit alpha-1/2	64.08	9	45	9	15.32	1.43
P68370	Tubulin alpha-1A chain	14.63	6	9	1	50.10	1.42
Q6AYZ1	Tubulin alpha-1C chain	10.69	5	6	1	49.91	1.42
Q68FR8	Tubulin alpha-3 chain	12.89	5	8	1	49.93	1.42
P16290	Phosphoglycerate mutase 2	8.70	2	3	2	28.74	1.42
Q91XR8	Phospholipid hydroperoxide glutathione peroxidase, nuclear	5.93	2	2	2	29.32	1.38
P36970	Phospholipid hydroperoxide glutathione peroxidase, mitochondrial	7.61	2	2	2	22.21	1.38
Q63041	Alpha-1-macroglobulin	0.93	2	2	2	167.02	1.35
P05508	NADH-ubiquinone oxidoreductase chain 4	8.28	4	7	4	51.77	1.34

P48675	Desmin	39.87	19	49	15	53.42	1.33
P00406	Cytochrome c oxidase subunit 2	22.03	5	25	5	25.93	1.32
P07895	Superoxide dismutase [Mn], mitochondrial	6.31	2	4	2	24.66	1.32
Q6P6R2	Dihydrolipoyl dehydrogenase, mitochondrial	8.84	3	4	3	54.00	1.30
P07335	Creatine kinase B-type	7.87	4	8	2	42.70	1.30
P17209	Myosin light chain 4	17.10	3	8	2	21.27	1.30
P84817	Mitochondrial fission 1 protein	17.11	3	6	3	16.98	1.30
B2RZ37	Receptor expression-enhancing protein 5	13.76	3	6	3	21.42	1.29
O88813	Long-chain-fatty-acid--CoA ligase 5	4.10	3	11	3	76.36	1.26
P80431	Cytochrome c oxidase subunit 7B, mitochondrial	10.00	2	4	2	8.99	1.26
P27274	CD59 glycoprotein	16.67	2	3	2	13.78	1.25
P53534	Glycogen phosphorylase, brain form (Fragment)	7.64	7	8	4	96.11	1.24
P60711	Actin, cytoplasmic 1	43.73	21	363	5	41.71	1.24
P18886	Carnitine O-palmitoyltransferase 2, mitochondrial	24.62	18	36	18	74.06	1.24
Q63619	Ubiquinone biosynthesis protein COQ7 homolog (Fragment)	11.17	2	3	2	20.13	1.24
P47819	Glial fibrillary acidic protein	3.49	2	2	1	49.93	1.23
P12346	Serotransferrin	8.02	6	8	6	76.35	1.23
Q9Z0V6	Thioredoxin-dependent peroxide reductase, mitochondrial	21.40	5	11	5	28.28	1.23
Q9JK11	Reticulon-4	1.89	2	3	2	126.31	1.23

P26284	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	22.82	10	26	10	43.20	1.23
Q3KR86	MICOS complex subunit Mic60 (Fragment)	32.68	19	40	19	67.14	1.23
P10888	Cytochrome c oxidase subunit 4 isoform 1, mitochondrial	48.52	10	100	10	19.50	1.22
P23928	Alpha-crystallin B chain	32.57	4	9	4	20.08	1.22
Q9JLJ3	4-trimethylaminobutyraldehyde dehydrogenase	3.64	2	2	2	53.62	1.21
P41350	Caveolin-1	14.61	4	5	2	20.54	1.21
P31000	Vimentin	12.23	7	13	2	53.70	1.21
Q68FU7	Ubiquinone biosynthesis monooxygenase COQ6	4.83	2	2	2	51.46	1.20
P00507	Aspartate aminotransferase, mitochondrial	41.16	18	47	18	47.28	1.20
P38718	Mitochondrial pyruvate carrier 2	33.86	5	10	5	14.25	1.20
P35704	Peroxiredoxin-2	18.69	4	5	4	21.77	1.19
P29117	Peptidyl-prolyl cis-trans isomerase F, mitochondrial	6.80	2	3	2	21.80	1.19
P07340	Sodium/potassium-transporting ATPase subunit beta-1	27.63	8	28	8	35.18	1.19
P00564	Creatine kinase M-type	28.87	12	31	10	43.02	1.19
P15650	Long-chain specific acyl-CoA dehydrogenase, mitochondrial	35.12	15	68	15	47.84	1.19
P16036	Phosphate carrier protein, mitochondrial	29.78	12	57	12	39.42	1.18
Q5XIT9	Methylcrotonoyl-CoA carboxylase beta chain, mitochondrial	6.75	4	5	4	61.48	1.18
P07943	Aldose reductase	9.81	3	3	3	35.77	1.18

P81155	Voltage-dependent anion-selective channel protein 2	60.00	14	42	12	31.73	1.17
Q07969	Platelet glycoprotein 4	17.80	9	18	9	52.70	1.17
P21396	Amine oxidase [flavin-containing]	21.29	13	16	13	59.47	1.17
P30427	Plectin	0.30	2	2	2	533.21	1.17
Q62651	Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial	7.34	3	4	3	36.15	1.17
P24368	Peptidyl-prolyl cis-trans isomerase B	25.00	5	6	5	23.79	1.17
P62632	Elongation factor 1-alpha 2	13.17	7	11	3	50.42	1.17
Q5XIF6	Tubulin alpha-4A chain	16.74	7	12	2	49.89	1.17
Q80W89	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 11	19.15	2	2	2	14.84	1.17
P62329	Thymosin beta-4	29.55	2	3	2	5.05	1.16
P63312	Thymosin beta-10	29.55	2	3	2	5.02	1.16
D3ZTX0	Transmembrane emp24 domain-containing protein 7	7.52	2	2	2	25.46	1.16
P06685	Sodium/potassium-transporting ATPase subunit alpha-1	33.43	31	91	19	112.98	1.16
P14046	Alpha-1-inhibitor 3	1.35	2	2	2	163.67	1.16
Q03626	Murinoglobulin-1	1.34	2	2	2	165.22	1.16
P21913	Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial	53.19	17	51	17	31.81	1.16
P05708	Hexokinase-1	17.21	17	31	14	102.34	1.15
Q01129	Decorin	11.86	5	6	4	39.78	1.15
Q920L2	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial	40.40	20	66	20	71.57	1.15
O88989	Malate dehydrogenase, cytoplasmic	27.54	9	19	9	36.46	1.15

Q9R1Z0	Voltage-dependent anion-selective channel protein 3	50.53	12	23	10	30.78	1.15
P40241	CD9 antigen	7.52	2	2	2	25.20	1.15
P97700	Mitochondrial 2-oxoglutarate/malate carrier protein	31.85	10	42	10	34.22	1.15
P19643	Amine oxidase [flavin-containing] B	2.88	2	3	2	58.42	1.15
P70470	Acyl-protein thioesterase 1	11.30	3	3	3	24.69	1.15
P18163	Long-chain-fatty-acid--CoA ligase 1	46.78	30	95	30	78.13	1.15
P47853	Biglycan	5.69	3	4	2	41.68	1.15
P62630	Elongation factor 1-alpha 1	9.09	5	8	1	50.08	1.15
Q68FT1	Ubiquinone biosynthesis protein COQ9, mitochondrial	25.00	6	12	6	35.12	1.14
P26772	10 kDa heat shock protein, mitochondrial	57.84	7	28	7	10.90	1.14
Q63716	Peroxiredoxin-1	22.11	5	7	5	22.10	1.14
Q68FU3	Electron transfer flavoprotein subunit beta	52.55	13	33	13	27.67	1.14
P48500	Triosephosphate isomerase	22.89	6	12	6	26.83	1.14
O35854	Branched-chain-amino-acid aminotransferase, mitochondrial	7.12	3	6	3	44.25	1.14
P21588	5'-nucleotidase	2.08	2	2	2	63.93	1.14
P07483	Fatty acid-binding protein, heart	57.89	8	24	8	14.77	1.14
P08461	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial	20.57	11	23	11	67.12	1.14
P08733	Myosin regulatory light chain 2, ventricular/cardiac muscle isoform	87.35	14	93	14	18.87	1.14
Q62812	Myosin-9	8.36	16	22	10	226.20	1.13

P80432	Cytochrome c oxidase subunit 7C, mitochondrial	28.57	2	11	2	7.37	1.13
Q99ML5	Prenylcysteine oxidase	5.16	2	2	2	56.25	1.13
P10111	Peptidyl-prolyl cis-trans isomerase A	12.20	3	5	3	17.86	1.13
Q01728	Sodium/calcium exchanger 1	1.75	2	2	2	108.12	1.13
D3ZHA0	Filamin-C	1.32	4	6	4	290.80	1.13
B0K020	CDGSH iron-sulfur domain-containing protein 1	34.26	3	20	3	12.09	1.13
P11608	ATP synthase protein 8	46.27	3	14	3	7.64	1.12
Q05962	ADP/ATP translocase 1	47.65	15	133	6	32.97	1.12
P70623	Fatty acid-binding protein, adipocyte	21.97	3	6	3	14.70	1.12
Q6PCT8	Succinate dehydrogenase [ubiquinone] cytochrome b small subunit, mitochondrial	10.69	2	2	2	16.97	1.12
Q2TA68	Dynamin-like 120 kDa protein, mitochondrial	7.50	8	9	8	111.24	1.12
P11951	Cytochrome c oxidase subunit 6C-2	46.05	4	17	4	8.45	1.11
P11232	Thioredoxin	20.95	2	2	2	11.67	1.11
P12075	Cytochrome c oxidase subunit 5B, mitochondrial	36.43	7	83	7	13.91	1.11
O88767	Protein deglycase DJ-1	23.81	5	7	5	19.96	1.11
P13803	Electron transfer flavoprotein subunit alpha, mitochondrial	41.74	12	31	12	34.93	1.11
P62074	Mitochondrial import inner membrane translocase subunit Tim10	14.44	2	3	2	10.33	1.11
P97521	Mitochondrial carnitine/acylcarnitine carrier protein	36.54	11	27	11	33.13	1.11
P19234	NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial	61.29	13	68	13	27.36	1.11

B5DEH2	Erlin-2	5.90	3	5	3	37.69	1.11
A2VD12	Pre-B-cell leukemia transcription factor-interacting protein 1	3.60	3	4	3	80.23	1.11
P68136	Actin, alpha skeletal muscle	69.50	26	445	2	42.02	1.10
P35435	ATP synthase subunit gamma, mitochondrial	44.69	12	91	12	30.17	1.10
P11884	Aldehyde dehydrogenase, mitochondrial	30.06	14	29	14	56.45	1.10
Q8VBU2	Protein NDRG2	4.31	2	4	2	40.75	1.10
Q6UPE1	Electron transfer flavoprotein-ubiquinone oxidoreductase, mitochondrial	44.97	24	95	24	68.16	1.10
P62076	Mitochondrial import inner membrane translocase subunit Tim13	36.84	3	6	3	10.45	1.10
Q5RKI0	WD repeat-containing protein 1	4.13	3	3	3	66.14	1.09
Q5XIN6	LETM1 and EF-hand domain-containing protein 1, mitochondrial	4.33	3	3	3	83.01	1.09
P29266	3-hydroxyisobutyrate dehydrogenase, mitochondrial	8.66	2	4	2	35.28	1.09
Q64119	Myosin light polypeptide 6	30.46	5	45	2	16.96	1.09
P11661	NADH-ubiquinone oxidoreductase chain 5	6.07	3	8	3	68.92	1.09
Q924S5	Lon protease homolog, mitochondrial	2.63	3	4	3	105.73	1.08
Q63610	Tropomyosin alpha-3 chain	18.95	8	27	2	28.99	1.08
Q6AXV4	Sorting and assembly machinery component 50 homolog	22.39	10	18	10	51.93	1.08
Q9QWN8	Spectrin beta chain, non-erythrocytic 2	0.63	2	2	2	270.90	1.08
P35286	Ras-related protein Rab-13	8.87	2	5	1	22.89	1.08

Q9Z0V5	Peroxiredoxin-4	6.96	2	2	2	30.99	1.08
Q03346	Mitochondrial-processing peptidase subunit beta	3.48	2	4	1	54.23	1.08
Q8VHF5	Citrate synthase, mitochondrial	31.55	14	73	14	51.83	1.07
D3ZAF6	ATP synthase subunit f, mitochondrial	30.68	3	19	3	10.45	1.07
P45953	Very long-chain specific acyl-CoA dehydrogenase, mitochondrial	40.92	27	76	27	70.71	1.07
P02770	Serum albumin	6.74	5	37	5	68.69	1.07
P35434	ATP synthase subunit delta, mitochondrial	13.69	2	11	2	17.58	1.07
O70351	3-hydroxyacyl-CoA dehydrogenase type-2	35.63	7	17	7	27.23	1.07
Q99376	Transferrin receptor protein 1 (Fragment)	5.47	4	4	4	70.11	1.07
Q9ER34	Aconitate hydratase, mitochondrial	45.26	32	117	32	85.38	1.07
Q5I0C3	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial	3.22	2	3	2	79.28	1.07
P20788	Cytochrome b-c1 complex subunit Rieske, mitochondrial	48.18	11	64	11	29.43	1.07
Q9QZ76	Myoglobin	56.49	7	43	7	17.15	1.06
P51868	Calsequestrin-2	16.22	7	18	7	47.81	1.06
P35280	Ras-related protein Rab-8A	14.01	3	8	1	23.65	1.06
Q7TQ16	Cytochrome b-c1 complex subunit 8	18.29	2	5	2	9.84	1.06
Q6NYB7	Ras-related protein Rab-1A	25.85	6	11	4	22.66	1.06
P08010	Glutathione S-transferase Mu 2	8.26	2	2	2	25.69	1.06
P61589	Transforming protein RhoA	9.33	2	2	2	21.77	1.06
P11980	Pyruvate kinase PKM	24.11	13	18	13	57.78	1.06
P45592	Cofilin-1	12.65	3	3	2	18.52	1.06

P62909	40S ribosomal protein S3	8.64	2	3	2	26.66	1.06
Q9ESS6	Basal cell adhesion molecule	7.37	4	4	4	67.47	1.06
P49432	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	34.54	11	32	11	38.96	1.05
P17764	Acetyl-CoA acetyltransferase, mitochondrial	40.09	16	51	16	44.67	1.05
Q64591	2,4-dienoyl-CoA reductase, mitochondrial	8.66	3	7	3	36.11	1.05
Q9WVA1	Mitochondrial import inner membrane translocase subunit Tim8 A	22.68	2	2	2	11.04	1.05
P04642	L-lactate dehydrogenase A chain	36.14	13	25	11	36.43	1.05
P56571	ES1 protein homolog, mitochondrial	32.71	8	11	8	28.16	1.05
P11240	Cytochrome c oxidase subunit 5A, mitochondrial	30.82	7	105	7	16.12	1.05
P10536	Ras-related protein Rab-1B	22.39	5	10	3	22.15	1.05
Q5HZA9	Transmembrane protein 126A	14.80	3	5	3	21.64	1.05
Q06647	ATP synthase subunit O, mitochondrial	60.09	12	71	12	23.38	1.05
P35171	Cytochrome c oxidase subunit 7A2, mitochondrial	27.71	2	13	2	9.35	1.05
Q01205	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial	22.25	9	17	9	48.89	1.05
Q6RJR6	Reticulon-3	2.87	3	4	3	101.45	1.05
P10719	ATP synthase subunit beta, mitochondrial	52.55	22	224	22	56.32	1.04

D4A3K5	Histone H1.1	7.48	2	3	2	21.99	1.04
P41565	Isocitrate dehydrogenase [NAD] subunit gamma 1, mitochondrial	7.12	4	7	3	42.82	1.04
P15865	Histone H1.4	17.81	4	9	4	21.97	1.04
P58775	Tropomyosin beta chain	23.24	11	33	5	32.82	1.04
Q29RW1	Myosin-4	16.30	40	323	1	222.74	1.04
P07632	Superoxide dismutase [Cu-Zn]	16.23	3	3	3	15.90	1.04
P63039	60 kDa heat shock protein, mitochondrial	34.03	17	40	17	60.92	1.04
P85972	Vinculin	15.20	17	19	17	116.54	1.04
P63269	Actin, gamma-enteric smooth muscle	65.96	24	448	2	41.85	1.03
P25809	Creatine kinase U-type, mitochondrial	8.37	4	16	3	47.00	1.03
Q4G069	Regulator of microtubule dynamics protein 1	12.90	5	8	5	35.38	1.03
P49134	Integrin beta-1	11.51	7	8	7	88.44	1.03
P68035	Actin, alpha cardiac muscle 1	69.50	26	463	2	41.99	1.03
P62738	Actin, aortic smooth muscle	69.50	26	463	2	41.98	1.03
P21571	ATP synthase-coupling factor 6, mitochondrial	50.93	6	19	6	12.49	1.03
P07633	Propionyl-CoA carboxylase beta chain, mitochondrial	11.28	5	7	5	58.59	1.03
P08503	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	22.09	9	23	9	46.53	1.03
P24473	Glutathione S-transferase kappa 1	11.50	3	3	3	25.48	1.02
Q63258	Integrin alpha-7	6.08	7	8	7	124.12	1.02
P06761	78 kDa glucose-regulated protein	35.32	22	37	19	72.30	1.02
P42123	L-lactate dehydrogenase B chain	34.13	12	28	10	36.59	1.02
P16409	Myosin light chain 3	81.50	16	211	12	22.14	1.02

P11517	Hemoglobin subunit beta-2	67.35	10	85	2	15.97	1.02
Q641Y2	NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial	38.23	13	61	13	52.53	1.02
P85834	Elongation factor Tu, mitochondrial	30.75	14	22	14	49.49	1.02
P04692	Tropomyosin alpha-1 chain	69.72	30	88	24	32.66	1.02
P29975	Aquaporin-1	5.95	2	3	2	28.84	1.02
P14668	Annexin A5	12.85	5	6	4	35.72	1.01
P23965	Enoyl-CoA delta isomerase 1, mitochondrial	32.87	10	17	10	32.23	1.01
P27952	40S ribosomal protein S2	7.85	3	3	3	31.21	1.01
Q9WVK7	Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial	18.79	9	20	9	34.43	1.01
P15651	Short-chain specific acyl-CoA dehydrogenase, mitochondrial	33.98	11	22	11	44.74	1.01
Q5U316	Ras-related protein Rab-35	10.95	2	6	0	23.01	1.01
P35289	Ras-related protein Rab-15	10.38	2	6	0	24.27	1.01
Q5XIH7	Prohibitin-2	52.17	18	35	18	33.29	1.01
P35281	Ras-related protein Rab-10	18.00	4	8	2	22.84	1.01
P67779	Prohibitin	49.26	12	29	12	29.80	1.01
Q68FX0	Isocitrate dehydrogenase [NAD] subunit beta, mitochondrial	7.79	4	8	3	42.33	1.01
Q8VID1	Dehydrogenase/reductase SDR family member 4	20.43	5	10	5	29.80	1.01
P62982	Ubiquitin-40S ribosomal protein S27a	9.62	2	5	2	17.94	1.01
P0CG51	Polyubiquitin-B	19.67	2	5	2	34.35	1.01
P62986	Ubiquitin-60S ribosomal protein L40	11.72	2	5	2	14.72	1.01
Q63429	Polyubiquitin-C	18.52	2	5	2	91.03	1.01

Q5BK63	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial	35.54	12	30	12	42.53	1.01
P54290	Voltage-dependent calcium channel subunit alpha-2/delta-1	2.20	2	3	2	123.75	1.01
P24329	Thiosulfate sulfurtransferase	5.39	2	3	2	33.39	1.00
Q09073	ADP/ATP translocase 2	42.28	15	86	6	32.88	1.00
Q66HF1	NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial	53.23	32	117	32	79.36	1.00
P48284	Carbonic anhydrase 4	7.77	2	2	2	35.05	1.00
P04797	Glyceraldehyde-3-phosphate dehydrogenase	31.23	10	23	10	35.81	1.00
P16086	Spectrin alpha chain, non-erythrocytic 1	0.57	2	2	2	284.46	1.00
P62260	14-3-3 protein epsilon	15.29	4	5	3	29.16	1.00
Q64536	[Pyruvate dehydrogenase (acetyl-transferring)] kinase isozyme 2, mitochondrial	3.69	2	2	2	46.08	1.00
Q02253	Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	28.22	14	30	14	57.77	1.00
P09605	Creatine kinase S-type, mitochondrial	37.95	15	110	14	47.36	1.00
P14669	Annexin A3	7.10	3	3	3	36.34	0.99
Q6AY30	Saccharopine dehydrogenase-like oxidoreductase	6.76	3	3	3	47.06	0.99
Q68FT3	Pyridine nucleotide-disulfide oxidoreductase domain-containing protein 2	7.57	5	6	5	62.84	0.99
P13221	Aspartate aminotransferase, cytoplasmic	7.99	4	5	4	46.40	0.99

Q60587	Trifunctional enzyme subunit beta, mitochondrial	33.05	16	54	16	51.38	0.99
Q9Z2L0	Voltage-dependent anion-selective channel protein 1	69.96	18	71	17	30.74	0.99
P29410	Adenylate kinase 2, mitochondrial	15.48	4	8	4	26.36	0.99
Q5XIF3	NADH dehydrogenase [ubiquinone] iron-sulfur protein 4, mitochondrial	31.43	5	16	5	19.73	0.99
P31399	ATP synthase subunit d, mitochondrial	69.57	10	35	10	18.75	0.98
P32551	Cytochrome b-c1 complex subunit 2, mitochondrial	44.69	19	90	19	48.37	0.98
A0JPQ4	Tripartite motif-containing protein 72	20.75	9	12	9	52.80	0.98
Q4V8F9	Hydroxysteroid dehydrogenase-like protein 2	11.83	7	13	7	58.31	0.98
P05065	Fructose-bisphosphate aldolase A	37.09	13	29	13	39.33	0.98
P70567	Tropomodulin-1	4.18	2	3	2	40.46	0.98
P48037	Annexin A6	21.55	16	26	15	75.71	0.98
P48721	Stress-70 protein, mitochondrial	40.50	26	45	25	73.81	0.98
P15999	ATP synthase subunit alpha, mitochondrial	61.48	36	220	36	59.72	0.97
P06686	Sodium/potassium-transporting ATPase subunit alpha-2	20.39	18	56	6	112.15	0.97
P14408	Fumarate hydratase, mitochondrial	29.78	15	42	15	54.43	0.97
Q66HD0	Endoplasmin	12.06	10	12	9	92.71	0.97
Q5XI32	F-actin-capping protein subunit beta	18.01	6	7	6	30.61	0.97
Q5XIJ4	Protein FAM210A	11.72	3	6	3	31.46	0.97
P27881	Hexokinase-2	3.49	4	7	1	102.48	0.97
Q9R063	Peroxiredoxin-5, mitochondrial	27.23	5	7	5	22.17	0.96
Q9Z2S9	Flotillin-2	11.92	6	7	6	47.01	0.96

Q63862	Myosin-11 (Fragments)	7.39	9	31	1	152.40	0.96
P56741	Myosin-binding protein C, cardiac-type	30.30	35	75	35	140.67	0.96
Q5XIC0	Enoyl-CoA delta isomerase 2, mitochondrial	15.60	5	8	5	42.99	0.96
Q5M9I5	Cytochrome b-c1 complex subunit 6, mitochondrial	60.67	5	17	5	10.42	0.95
P38983	40S ribosomal protein SA	7.12	2	3	2	32.80	0.95
P07323	Gamma-enolase	8.53	3	5	2	47.11	0.95
P02564	Myosin-7	51.68	114	931	16	222.95	0.95
P29457	Serpin H1	7.67	4	6	4	46.49	0.95
P63018	Heat shock cognate 71 kDa protein	29.72	18	26	14	70.83	0.95
P11598	Protein disulfide-isomerase A3	20.59	11	16	11	56.59	0.95
P14659	Heat shock-related 70 kDa protein 2	13.27	8	10	4	69.60	0.95
P15800	Laminin subunit beta-2	6.89	12	13	12	196.35	0.95
P52504	NADH dehydrogenase [ubiquinone] iron-sulfur protein 6, mitochondrial	18.97	2	5	2	12.78	0.95
P62898	Cytochrome c, somatic	39.05	5	18	5	11.60	0.94
Q07439	Heat shock 70 kDa protein 1A/1B	5.93	4	6	2	70.14	0.94
P55063	Heat shock 70 kDa protein 1-like	8.42	5	7	2	70.51	0.94
P13437	3-ketoacyl-CoA thiolase, mitochondrial	54.66	16	39	16	41.84	0.94
P30839	Fatty aldehyde dehydrogenase	3.31	2	2	2	54.05	0.94
Q499N5	Acyl-CoA synthetase family member 2, mitochondrial	16.91	11	17	11	67.84	0.94
P0C219	Sarcolemmal membrane-associated protein	1.98	2	2	2	98.16	0.94
P29418	ATP synthase subunit epsilon, mitochondrial	45.10	3	9	3	5.76	0.94

P18418	Calreticulin	18.99	9	20	9	47.97	0.94
P19511	ATP synthase F(0) complex subunit B1, mitochondrial	42.58	13	51	13	28.85	0.94
Q9EQP5	Prolargin	4.24	2	3	2	43.15	0.94
B0LPN4	Ryanodine receptor 2	3.59	18	23	18	562.59	0.94
P36201	Cysteine-rich protein 2	8.65	2	2	2	22.68	0.94
P56574	Isocitrate dehydrogenase [NADP], mitochondrial	44.25	20	79	20	50.94	0.94
Q9Z1P2	Alpha-actinin-1	10.99	9	24	4	102.90	0.93
Q64578	Sarcoplasmic/endoplasmic reticulum calcium ATPase 1	15.29	15	77	15	109.34	0.93
P07153	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	3.31	2	2	2	68.26	0.93
Q64428	Trifunctional enzyme subunit alpha, mitochondrial	47.31	31	114	31	82.61	0.93
P02563	Myosin-6	61.40	146	1232	43	223.37	0.93
P34058	Heat shock protein HSP 90-beta	12.29	9	13	5	83.23	0.93
P10860	Glutamate dehydrogenase 1, mitochondrial	15.95	9	15	9	61.38	0.92
B2GV06	Succinyl-CoA:3-ketoacid coenzyme A transferase 1, mitochondrial	21.92	9	20	9	56.17	0.92
Q4V8H8	EH domain-containing protein 2	11.97	7	11	6	61.20	0.92
Q561S0	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 10, mitochondrial	26.48	9	37	9	40.47	0.92
Q07936	Annexin A2	11.50	4	4	4	38.65	0.92
P05197	Elongation factor 2	3.26	3	5	3	95.22	0.92
P63219	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-5	23.53	2	2	2	7.31	0.92

P09811	Glycogen phosphorylase, liver form	3.06	3	3	1	97.42	0.92
P38652	Phosphoglucomutase-1	6.41	3	3	3	61.37	0.91
P18596	Sarcoplasmic/endoplasmic reticulum calcium ATPase 3	7.26	7	34	7	116.21	0.91
Q6P7S1	Acid ceramidase	4.82	2	2	2	44.42	0.91
Q5XI78	2-oxoglutarate dehydrogenase, mitochondrial	18.28	18	41	18	116.22	0.91
P16617	Phosphoglycerate kinase 1	18.94	9	17	9	44.51	0.91
P26453	Basigin	15.46	6	13	6	42.41	0.91
Q6RUV5	Ras-related C3 botulinum toxin substrate 1	18.23	4	5	4	21.44	0.91
P24268	Cathepsin D	19.41	7	10	7	44.65	0.90
Q9ESV6	Glyceraldehyde-3-phosphate dehydrogenase, testis-specific	4.86	2	6	2	46.68	0.90
P11507	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	42.28	41	170	41	114.69	0.90
Q9JM53	Apoptosis-inducing factor 1, mitochondrial	11.44	7	11	7	66.68	0.90
P12007	Isovaleryl-CoA dehydrogenase, mitochondrial	19.81	9	14	9	46.41	0.90
P14882	Propionyl-CoA carboxylase alpha chain, mitochondrial	15.74	11	16	11	81.57	0.90
Q63704	Carnitine O-palmitoyltransferase 1, muscle isoform	16.06	12	24	12	88.16	0.90
Q4QQV3	Protein FAM162A	15.48	2	3	2	17.82	0.90
P04764	Alpha-enolase	26.27	10	22	7	47.10	0.90
P13086	Succinyl-CoA ligase [ADP/GDP-forming] subunit alpha, mitochondrial	17.63	6	9	6	36.13	0.90
Q63081	Protein disulfide-isomerase A6	10.23	4	5	4	48.14	0.89

Q5U2X7	Mitochondrial import inner membrane translocase subunit Tim21	16.73	4	5	4	27.86	0.89
O88267	Acyl-coenzyme A thioesterase 1	6.68	4	7	4	45.98	0.89
Q68FY0	Cytochrome b-c1 complex subunit 1, mitochondrial	30.00	15	70	14	52.82	0.89
Q3T1K5	F-actin-capping protein subunit alpha-2 OS=Rattus norvegicus GN=Capza2 PE=1 SV=1	15.03	3	5	3	32.95	0.89
P05712	Ras-related protein Rab-2A	14.15	2	2	2	23.52	0.89
Q99NA5	Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial	16.94	7	13	7	39.59	0.89
O35763	Moesin	2.77	2	2	2	67.70	0.89
P85125	Polymerase I and transcript release factor	18.37	5	8	5	43.88	0.89
Q03410	Synaptonemal complex protein 1	0.70	2	4	2	116.44	0.88
P09495	Tropomyosin alpha-4 chain	16.94	7	16	1	28.49	0.88
Q63355	Unconventional myosin-Ic	3.54	4	5	4	119.74	0.88
P13832	Myosin regulatory light chain RLC-A	16.28	3	5	3	19.88	0.87
Q5I0E7	Transmembrane emp24 domain-containing protein 9	11.06	4	5	4	27.01	0.87
P09812	Glycogen phosphorylase, muscle form	11.16	12	14	9	97.21	0.87
P61983	14-3-3 protein gamma	6.88	2	3	1	28.29	0.87
O55171	Acyl-coenzyme A thioesterase 2, mitochondrial	11.04	6	10	6	49.67	0.87
Q6IFU8	Keratin, type I cytoskeletal 17	6.24	3	5	1	48.09	0.86
Q641Z6	EH domain-containing protein 1	4.68	3	5	2	60.57	0.86
Q5SGE0	Leucine-rich PPR motif-containing protein, mitochondrial	11.35	17	21	17	156.55	0.86

P97576	GrpE protein homolog 1, mitochondrial	15.21	4	6	4	24.28	0.86
P15429	Beta-enolase	20.74	9	16	6	46.98	0.86
P29419	ATP synthase subunit e, mitochondrial	71.83	5	17	5	8.25	0.85
P29147	D-beta-hydroxybutyrate dehydrogenase, mitochondrial	13.12	4	6	4	38.18	0.85
P35565	Calnexin	7.78	5	8	5	67.21	0.85
P63031	Mitochondrial pyruvate carrier 1	22.02	4	7	4	12.45	0.84
Q63362	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 5	14.66	3	8	3	13.40	0.84
Q66H98	Serum deprivation-response protein	7.91	3	3	3	46.36	0.84
Q62902	Protein ERGIC-53	2.71	2	2	2	57.92	0.83
P62836	Ras-related protein Rap-1A	15.76	3	6	3	20.97	0.82
Q62636	Ras-related protein Rap-1b	15.76	3	6	3	20.79	0.82
Q9QXQ0	Alpha-actinin-4	7.46	7	9	2	104.85	0.82
P11762	Galectin-1	40.74	6	9	6	14.85	0.82
P82995	Heat shock protein HSP 90-alpha	6.41	5	6	2	84.76	0.81
Q9QY17	Protein kinase C and casein kinase substrate in neurons 2 protein	3.28	2	2	2	55.94	0.80
Q6PDU7	ATP synthase subunit g, mitochondrial	26.21	3	10	3	11.43	0.79
Q704S8	Carnitine O-acetyltransferase	8.63	6	8	6	70.76	0.79
P04785	Protein disulfide-isomerase	24.17	13	18	13	56.92	0.79
P63102	14-3-3 protein zeta/delta	6.53	2	2	1	27.75	0.78
P62804	Histone H4	57.28	6	50	6	11.36	0.75
Q6VBQ5	Myeloid-associated differentiation marker	9.43	2	3	2	35.13	0.71
P0C0S7	Histone H2A.Z	20.31	3	34	1	13.55	0.71

P54708	Potassium-transporting ATPase alpha chain 2	4.73	5	11	1	114.90	0.70
Q7M0E3	Destrin	12.12	2	2	1	18.52	0.67
Q00715	Histone H2B type 1	35.20	5	24	1	13.98	0.66
P02262	Histone H2A type 1	32.31	4	37	2	14.07	0.63
A9UMV8	Histone H2A.J	32.56	4	37	2	14.04	0.63
P0C170	Histone H2A type 1-E	32.31	4	37	2	14.11	0.63
P84245	Histone H3.3	22.06	4	11	4	15.32	0.62
Q6LED0	Histone H3.1	29.41	6	13	6	15.39	0.62
Q00729	Histone H2B type 1-A	33.07	5	18	1	14.22	0.56

The proteomics data were further processed using the Cluster 3.0 software and IPA. A subset of 50 proteins, consisting of 12 sarcomeric proteins, 10 ROS-scavenging proteins, 10 apoptotic proteins, and 15 abundant proteins, underwent hierarchical clustering analysis (**Figure 3-6**). Differences in protein abundances were observed across all the LAD ligation/reperfusion and sham replicates, and hierarchical clustering revealed no obvious grouping among the replicates (**Figure 3-6**). Overall, the cluster analysis results illustrated the slight increase in whole proteome ROS-scavenging proteins protein abundance after LAD ligation/reperfusion as compared to sham surgery, and no observable difference in the remaining proteins investigated. IPA core analysis of the identified proteome proteins developed 17 protein interaction networks, with metabolic disease/disorder (score 52, 29 proteins), free radical scavenging (score 49, 28 proteins), and energy production/nucleic acid metabolism (score 37, 23 proteins) the highest scored pathways.

A comparison of the quantitative proteomics results for free cardiomyocyte proteins and proteins covalently trapped on DNA revealed a total of 78 proteins (86.7%) that were identified in both experiments. Of the 78 proteins shared between the two experiments, 14 proteins increased in abundance after LAD ligation/reperfusion compared to sham surgery in both whole protein and DPC samples, while the rest were only increased for the cross-linked proteins, but not the global proteome (**Figure 3-7**). Overall, these results indicate little correlation between global protein amounts and DPC formation, suggesting that oxidative cross-linking between proteins and DNA is induced by the influx of hydroxyl radicals during reperfusion of cardiac tissues.

Figure 3-6: Heat map analysis of the TMT abundance data for 47 representative proteins identified in the whole proteome. Proteins analyzed include sarcomere proteins (13), ROS scavenging proteins (11), apoptosis regulating proteins (7), and the remaining most abundant proteins (16).

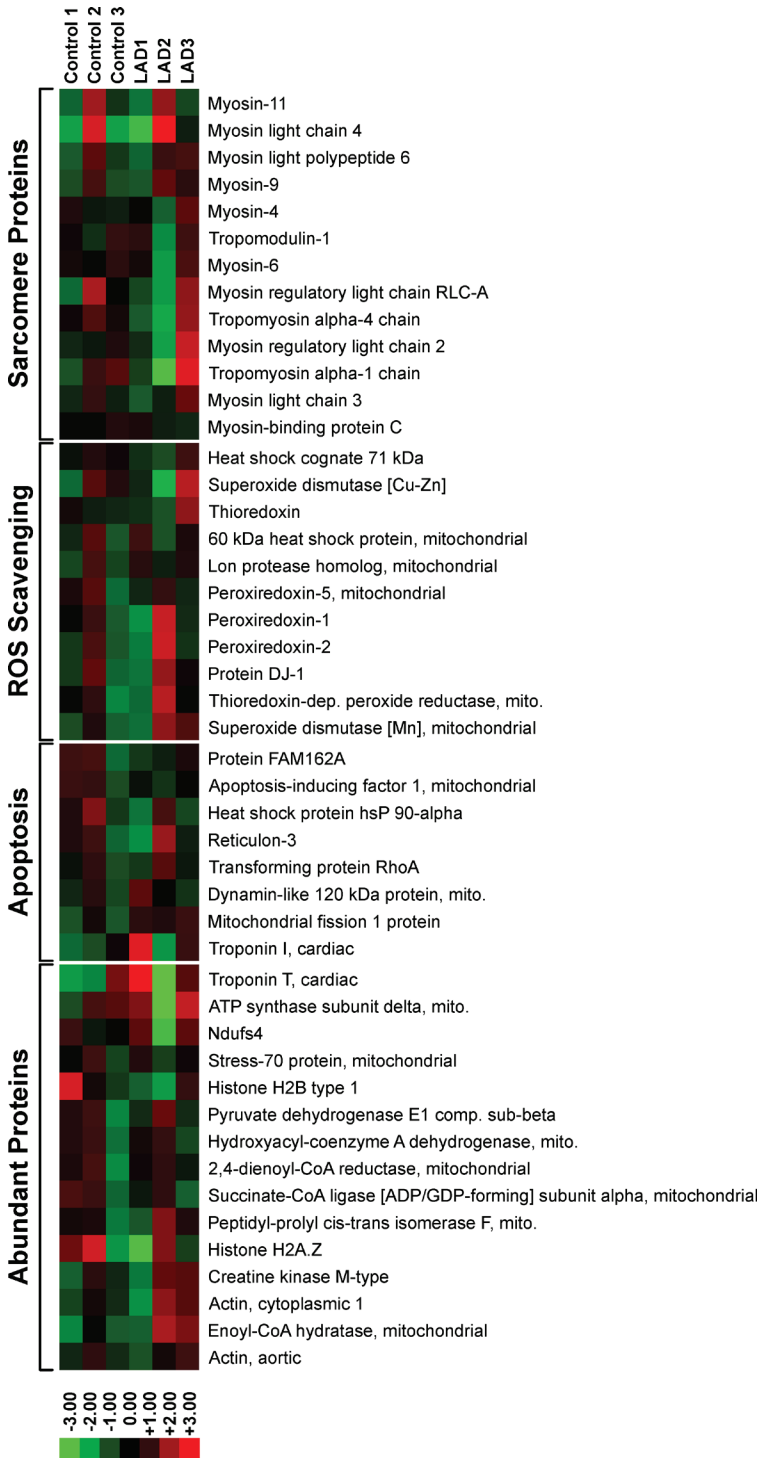
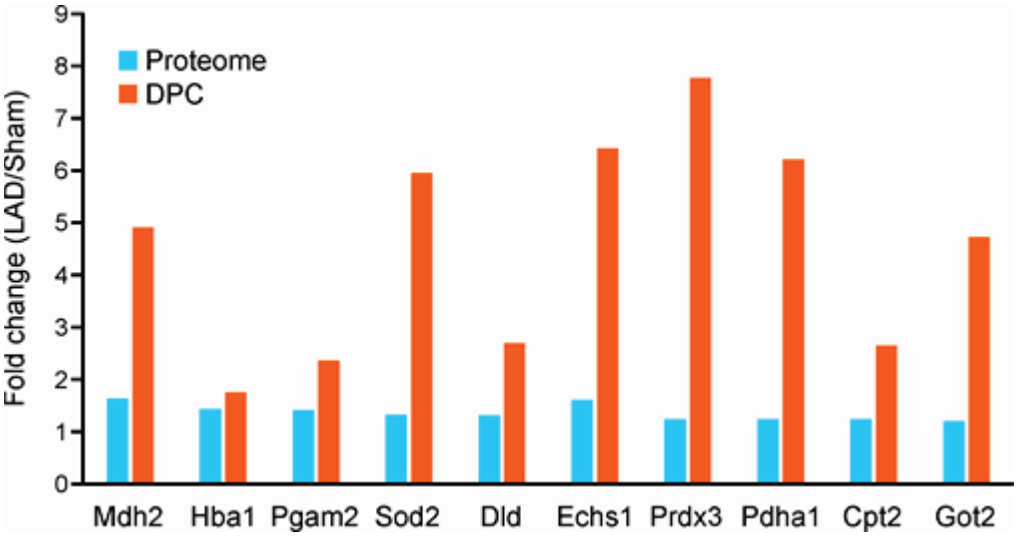


Figure 3-7: Changes in protein abundances in DPCs (red) *vs* whole proteome (blue). The protein abundances (LAD ligation/sham) in the whole proteome and DPCs of 10 representative proteins were compared. Protein names correspond to gene symbols: Mdh2, malate dehydrogenase, mitochondrial; Hba1, hemoglobin subunit-alpha ½; Pgam2, phosphoglycerate mutase 2; Sod2, superoxide dismutase [Mn], mitochondrial; Dld, dihydrolipoyl dehydrogenase, mitochondrial; Echs1, enoyl-CoA hydratase, mitochondrial; Prdx3, thioredoxin-dependent peroxide reductase, mitochondrial; Cpt2, carnitine O-palmitoyltransferase 2, mitochondrial; Got2, aspartate aminotransferase, mitochondrial.



3.4. Discussion

DNA-protein crosslinks are super-bulky DNA adducts capable of interfering with crucial biological functions such DNA replication and transcription.[262] Recent studies in our laboratory have shown that DPC lesions completely block human DNA polymerases.[265-267, 337] Similarly, Nakano *et al* reported that DPCs over 14.1 kDa in size located on the translocating strand completely block DNA unwinding by human helicases (mini chromosome maintenance Mcm467 subcomplex), while DPCs between 5 to 14.1 kDa severely inhibit helicase progression.[264] Recently, our laboratory found that bacterial RNA polymerases were blocked by protein lesions conjugated to the N7 position of guanine (Ji and Tretyakova, unpublished data).

Covalent DPCs readily form in cells via a radical mechanism in the presence of reactive oxygen species. Izzotti *et al* previously reported age-related DPC accumulation in the brain, liver, and heart of 12 or 24-month-old mice.[270] In mouse myocardium, DPCs accounted for 2.2 ± 0.3% of total DNA in 12-month old mice and 3.4 ± 0.5% of total DNA in 24-month old mice.[270] Furthermore, there was a strong correlation between the levels of measured 8-oxoguanine and DPCs ($r = 0.800$, $P<0.0001$), indicating the DPC formation was related to ROS exposure.[270] However, protein identities have remained unknown, and no absolute quantitation of DPCs was conducted.

Every year, 735,000 people in the United States suffer myocardial infarction.[364, 365] Although reperfusion surgery can successfully restore the blood flow to affected myocardium, a sudden influx of oxygenated blood leads to a spike in ROS and reperfusion injury, which is associated with acute cardiomyocyte necrosis, and potentially chronic heart disease.[366, 367, 486] An estimated 40-50% of the final MI damage has been attributed to

reperfusion injury.[367] However, the exact mechanism of ROS-mediated myocardium damage has yet to be elucidated. We hypothesized that the influx of hydroxyl radicals caused by ischemia/reperfusion leads to the formation of toxic DNA-protein cross-links, which may contribute to reperfusion injury.

In the present study, we employed a well-established animal model (LAD ligation/reperfusion) to investigate DPC formation in cardiomyocytes following ischemia-reperfusion injury.[487] Animals in treatment group that underwent LAD ligation/reperfusion exhibited greatly elevated plasma levels of cTnI (>50 ng/mL), confirming severe MI (**Table 3-1**). Furthermore, plasma levels of the oxidative stress biomarker 8-iso-PGF2α were increased 3-5-fold in animals that underwent the LAD surgery and reperfusion as compared to sham surgery (**Figure 3-1**), although pre-surgery 8-iso-PGF2α may have been elevated because anesthetization created a hyperoxic environment. Taken together, our results shown in **Table 3-1** and **Figure 3-1** confirm that animals in the LAD group experienced MI/reperfusion injury that led to an increase in oxidants production in the heart.

Two general approaches were employed to investigate DPC formation in rat cardiomyocytes following ischemia-reperfusion injury. Absolute numbers of hydroxyl radical-mediated DPCs were established using the sensitive and accurate isotope dilution nanoLC-ESI$^+$-MS/MS method developed in our laboratory (**Figure 3-2**). Protein identities were established using quantitative proteomics, which was coupled with analysis of the global proteome in order to detect any system-wide changes in protein levels. It was necessary to employ quantitative proteomics in this case because background levels of DPCs are found in the normal heart as a result of aerobic respiration.

In order to determine the absolute numbers of hydroxyl radical-induced DPC in cardiomyocytes following ischemia/reperfusion injury, DPC-containing DNA was digested with nucleases and proteinases to single monomer level, and the amounts of thymidine-tyrosine conjugates (dT-Tyr, **Scheme 3-2**) were determined by isotope dilution nanoLC-ESI^{+}-MS/MS using ^{13}C labeled internal standard synthesized in our laboratory (**Figure 3-2A**). dT-Tyr conjugate numbers were significantly increased in cardiomyocytes of rats that underwent ischemia-reperfusion injury as compared to sham controls (**Figure 3-2B**). Altman *et al* previously measured the structurally similar 3- [(1,3-dihydro-2,4-dioxopyrimidin-5- yl jmethyl] -L-tyrosine (Thy-Tyr) adduct in SP2/0 murine cells treated with Fe(II) using a gas chromatography mass spectrometry assay.[488] Treatment with 1 mM Fe(II) yielded 32 Thy-Tyr adducts per 10^5 nucleotides,[488] nearly a 1000-fold increase over our highest observed dT-Tyr levels. They were also able to detect up to 15 Thy-Tyr adducts per 10^6 nucleotides in control cells,[488] which are once again significantly higher than what weobserved in our sham surgery samples. Taken together, this indicates that both cell culture and treatment of cells with Fe(II) may expose cells to levels of hydroxyl radicals that are not physiologically relevant.[489]

Proteins covalently attached to DNA were identified and quantified using quantitative proteomics. These analyses have revealed a highly significant increase in radical-induced DPC formation following LAD ligation/reperfusion surgery as compared to sham surgery (**Table 3-2**, **Figure 3-4,** and **Figure 3-5**). Quantitative proteomic analysis of DNA-conjugated proteins has identified a total of 90 proteins: of these, 80 proteins were increased in abundance at least 1.2-fold after LAD ligation/ reperfusion as compared to

sham surgery and 1 protein was decreased in abundance under the same conditions (**Table 3-2**).

A large fraction of the oxidatively cross-linked proteins (43.3%) were mitochondrial in nature, including cytochrome C oxidase, pyruvate dehydrogenase, and enoyl-CoA hydratase (**Table 3-2**). Since our DPC extraction methodology does not discriminate chromosomal DNA from mitochondrial DNA, it is possible that these proteins are covalently trapped on mitochondrial DNA. Mitochondria increase the production of superoxide anions, hydrogen peroxide, and hydroxyl radicals in response to ischemia/reperfusion.[465, 466, 490, 491] For example, all four complexes of the electron transport chain can produce reactive superoxide anions, and ischemic damage to complex I and III is believed to enhance the production of superoxides released during reperfusion injury.[402, 466, 492] Furthermore, NAD-dependent mitochondrial pyruvate dehydrogenase (PDH), a protein identified in the DPC samples, can produce both O_2^- and H_2O_2 as enzymatic byproducts.[493, 494] The mitochondrial NADH/NAD^+ ratio is elevated post-ischemia, leading to an increased enzymatic activity NAD-dependent dehydrogenases such as PDH and an increase in O_2^- and H_2O_2 byproducts.[493] Therefore, it is not surprising that mitochondrial proteins are exposed to increased oxidant levels and thus susceptible to becoming covalently trapped on mitochondrial DNA.

Additionally, many of the identified DPC protein constituents are known to be present in both the mitochondria or cytoplasm and the nucleus. For example, pyruvate dehydrogenase E1 components, which increased in DPC abundance 9.2-fold after LAD ligation/reperfusion compared to sham surgery, are classified as mitochondrial proteins. However, mitochondrial stress stimulates the translocation of pyruvate dehydrogenases to

the nucleus to generate nuclear acetyl CoA, which is required for histone acetylation and the activation of transcription.[495]

Ironically, ROS scavenging enzymes, thioredoxin-dependent peroxide reductase, superoxide dismutase [Mn], superoxide dismutase [Cu-Zn], and catalase have been found to participate in cross-linking to DNA following reperfusion injury (**Table 3-2**). Although these ROS scavenging enzymes are induced in response to oxidative stress, superoxide dismutases and catalase are known to release hydroxyl radicals as a byproduct of peroxide metabolism.[466, 496, 497] Hydroxyl radicals are extremely reactive species capable of abstracting hydrogen ($H^{\bullet}$) from nearby DNA bases.[31, 35, 498] The resulting radical intermediates can react with nearby dismutase or catalase enzyme, yielding a DPC with the ROS scavenging protein.

The observed increase in DNA-protein cross-linking (**Figure 3-3**) cannot be attributed to changes in global protein levels following reperfusion injury. Only modest increases in protein biomarkers of cardiomyocyte injury after LAD ligation/reperfusion compared to sham surgery, while the majority of the proteome was unchanged (**Table 3-3**, **Figure 3-5,** and **Figure 3-6**). Increased levels of troponin-T have been established as a specific biomarker of myocardial damage and the breakdown of the contractile apparatus.[499-501] Furthermore, previous studies have observed the degradation and release of other sarcomere proteins including troponin-I, myosin light chain 1, and tropomyosins after LAD ligation/reperfusion in rats.[502] Our observation of increased levels of the sarcomere proteins troponin-T, troponin-I, myosin-3, myosin light chain 4, and actin after LAD ligation/reperfusion is also consistent with the breakdown of the contractile apparatus

after reperfusion injury, and may indicate damaged sarcomere proteins are translocated to the cytoplasm of cardiac cells as well as into the extracellular plasma.

Our results for global proteome changes are consistent with previous studies that utilized quantitative proteomics techniques to measure changes in protein expression after LAD ligation or LAD ligation/reperfusion injury compared to sham surgery.[501, 503] De Celle *et al* utilized two-dimensional gel electrophoresis quantification and mass spectrometry-based proteomics to identify 8 proteins (troponin T, heat shock protein 20 and 27, heterogeneous nuclear ribonucleoprotein K, pyruvate dehydrogenase E1-β, annexin A3, adenylate kinase 1, and serum amyloid p-component precursor) that were upregulated after LAD ligation/reperfusion and two proteins (myosin heavy chain α and catechol-o-methyltransferase) that were downregulated after LAD ligation/reperfusion.[501] Our quantitative proteomics analysis also observed upregulation of troponin T, pyruvate dehydrogenase E1-β, and annexin-3 after LAD ligation/reperfusion. Furthermore, we observed upregulation of two other heat shock proteins (10 kDa heat shock protein and 60 kDa heat shock protein mitochondrial) and adenylate kinase 2 (**Figure 3-6, Table 3-3**).

In another study, Liu *et al* utilized a label-free spectral counting analysis to identify 15 mitochondrial proteins that became upregulated after LAD ligation as compared to a sham surgery.[503] Our current study identified all 15 mitochondrial proteins, with six proteins (amine oxidase, mitochondrial 2-oxogluterate/malate carrier protein, ATP synthase subunit gamma, aldehyde dehydrogenase, dihydrolipoyllysine residue acetyltransferase component of pyruvate dehydrogenase complex, and electron transfer flavoprotein subunit alpha) upregulated after LAD ligation/reperfusion.[503] These changes

in mitochondrial protein expression are indicative of the ischemia-induced mitochondrial dysfunction and dysregulation of the electron transport chain.

3.5. Conclusions

To our knowledge, our study is the first to provide conclusive evidence for the increased DPC formation following ischemia/reperfusion injury. Our results suggest that oxidative DNA-protein cross-linking formation is a biological consequence of a myocardial infarction/ reperfusion. Furthermore, the quantitative nanoLC-ESI$^+$-MS/MS assay for DNA-protein cross-links reported here has the potential to improve the development of antioxidant and preventative reperfusion injury therapies. Previous experimental and clinical investigations of antioxidant therapies have reported mixed or inconsistent results.[504, 505] This partially may be caused by the lack of specificity and sensitivity of assays used to measure oxidant production.[506-508] Hydroxyl radical-induced dT-Tyr conjugates could serve as a novel biomarker of oxidant-induced tissue damage, providing a more accurate measure of the efficacy of future therapies designed to reduce oxidant levels brought by reperfusion.

4. Spartan Metalloprotease and Proteasomal Degradation in Repair of ROS-induced Chromosomal DNA-Protein Cross-Links in Cells and Tissues

4.1. Introduction

DNA-protein cross-links (DPCs) are super-bulky DNA lesions that form when proteins are irreversibly trapped on chromosomal DNA.[262, 263] DPCs can be induced after exposure to various physical and chemical agents including ionizing radiation,[274, 275] UV light,[276-278] transition metal ions such as Ni and Cr,[279-282] halogenated hydrocarbons,[283, 284] environmental carcinogens,[285] and common anticancer drugs such as nitrogen mustards and platinum compounds.[286-289] Endogenous DPCs can form when DNA-binding proteins such as topoisomerases,[290, 291] polymerases,[29, 30] and repair enzymes[509, 510] become trapped on DNA intermediates. Furthermore, endogenous DPCs are induced upon exposure to reactive oxygen species (ROS) produced during cellular respiration and inflammation,[292-294] and aldehyde byproducts of lipid peroxidation such as methylglyoxal[295] and malondialdehyde.[296] Due to their bulky nature, DPCs can interfere with DNA replication[264-267] and gene expression.[268] DPC formation is proposed to play a role in aging,[269, 270] cancer,[271, 272] cardiovascular disease,[269, 270] and neurodegenerative

disorders.[273] Despite the significant threat DPCs pose to cells, the mechanism(s) of DPC recognition and active DPC repair has yet to be fully elucidated.

Several canonical DNA repair pathways including nucleotide excision repair (NER) coupled to proteolysis of the protein constituent[263, 309, 426, 427] and homologous recombination/Fanconi anemia (HR) repair pathways[424, 425, 511] have been implicated in DPC repair. Stingele *et al.*[421] discovered a protease-dependent DPC repair mechanism in the yeast. These authors demonstrated the yeast metalloprotease Wss1 (weak suppressor of smt3) coordinates with tyrosyl-DNA-phosphodiesterase 1 or 2 (TPD1 and TPD2 respectively) to enzymatically process topoisomerase 1 and 2 cleavage complexes (Top1ccs and Top2ccs respectively) and allow for replication to proceed past topoisomerase-DNA cross-links.[421] It was further shown that Wss1 protects yeast cells from FA-induced toxicity.[421, 428]

In eukaryotic systems, proteolytic digestion of a DPC was first observed by Duxin *et al*.[430] They incubated a Dnmt1-containing plasmid and the corresponding plasmid control in *Xenopus* egg extract, followed by restriction enzyme digestion of the plasmids to yield a 165-nucleotide fragment containing the DPC.[430] Gel electrophoresis of the digestion lysates revealed one band for the control and multiple bands for the DPC-containing plasmid, indicating the protein constituent was proteolytically digested to smaller peptides.[430] Mammalian cells express the metalloprotease DVC1/Spartan (Spartan), which is considered the mammalian homolog of the yeast Wss1.[512] Spartan and Wss1 both share a similar domain organization and have a common evolutionary origin.[428, 512] Furthermore, Stingele *et al* have shown that isolated Spartan can proteolytically digest the DNA-binding proteins H1, H2A, H2B, H3, and Hmg1 in the presence of ssDNA, but

will only perform autocleavage in the presence of dsDNA or without DNA.[431] Due to this homology and confirmed activity, Spartan has been investigated as the protease required for DPC processing in mammalian cells.[430, 431, 433, 513]

Recently, Spartan-mediated digestion of a protein has been linked to the stalling of the replisome during DNA replication. For example, Morocz *et al* utilized a DNA fibre assay to demonstrate that Spartan is required for the immediate bypass of DPCs.[433] When Spartan-depleted cells were treated with FA to induce DPC formation, the length of newly formed DNA tracks was significantly decreased compared to FA-treated wild-type cells, indicating a strong inhibition of bypass across the FA-induced damage in the absence of Spartan.[433] When the DNA fibre assay was performed with cells expressing Spartan with a mutated protease domain (SprT), DNA replication was inhibited to the same degree as Spartan depletion.[433] These results are consistent with the replication-coupled DPC repair model proposed by Stingele *et al*[428, 431] and Vaz *et al*.[512, 513] During replication, the replisome/CMG helicase cannot bypass a DPC site and becomes stalled.[264, 421] This triggers Spartan-mediated proteolytic degradation of the protein constituent of DPC, and the resulting peptide-DNA adduct is immediately bypassed DNA polymerase ζ[430, 433, 512] and ultimately repaired by nucleotide excision repair.[309, 426, 427]

Although the DPC protease activity of Spartan has been established for the repair of Top1ccs and FA-induced DPCs, Spartan's involvement in removing other types of DPCs, such as ROS-induced DPCs, has not been demonstrated. Furthermore, Maskey *et al* were unable to detect differences in sensitivity of Spartan-deficient cells to etoposide (topoisomerase II DPCs), MK-4827 (PARP-DNA conjugates), and 5-aza-2'-deoxycytidine (DNA-methyltransferase DPCs), indicating that protein connectivity and protein identity

may impact Spartan's DPC-protease activity.[514] Finally, the molecular mechanism of Spartan recruitment and recognition of a DPC has yet to be fully elucidated.

Alternatively, proteolytic degradation of DPCs can be accomplished via proteasome-mediated degradation.[309, 427, 515-517] In mammalian systems, Top1ccs and Top2ccs are first proteolytically digested by the 26S proteasome to a smaller DNA-peptide cross-link,[515-518] which are then enzymatically cleaved from DNA by TPD1 and TPD2 respectively.[519-521] The first evidence of the 26S proteasome digesting non-topoisomerase DPCs was provided by Quievryn and Zecevic *et al*, who utilized the K^+/SDS precipitation method to demonstrate pharmacological inhibition of the 26S proteasome decreased the repair of FA and Cr(IV)-induced DPCs respectively.[309, 522] Furthermore, Baker *et al* utilized fluorescence-activated cell sorting (FACS) to demonstrate that inhibition of the 26S proteasome decreased the repair of a Dmnt-containing plasmid by 50% in Chinese hamster ovary (CHO) cells.[427] However, the relatively small decrease in repair observed by both Quievryn *et al* and Baker *et al* demonstrates that the 26S proteasome cannot independently digest the protein constituent of a DPC. Furthermore, no previous investigations have demonstrated that the 26S proteasome is involved in the repair of more ubiquitous DPCs, such as those induced by a radical-induced mechanism.

In the present study, we employed a combination of cytotoxicity screening and mass spectrometry based DPC quantitation to investigate the role of Spartan and the 26S proteasome in the repair of DPCs induced via a free radical mechanism. The common reactive oxygen species hydroxyl radicals can abstract a hydrogen from the methyl group of a thymidine in DNA, which can undergo an one electron addition to the 3-position of tyrosine residue in proteins to form a cross-link.[362, 363, 481, 482] We recently developed a

sensitive nanoLC-ESI$^+$-MS/MS assay for the total digestion product 2-amino-3-(4-hydroxy-3-((1-((2R,4R,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)methyl)phenyl)propanoic acid (dT-Tyr) to perform absolute quantitation of hydroxyl radical-induced DPC formation in cardiomyocytes (**Chapter 3**). Ionizing radiation (IR), a high energy radiation that induces the release of electrons from molecules, also generates hydroxyl radicals capable of inducing DPCs via the same mechanism.[523, 524] We employed cell viability assays to investigate the effects of exposing Spartan-deficient and 26S proteasome-inhibited cells to DPC inducing agents. Furthermore, quantitative HPLC-ESI$^+$-MS/MS methodologies were employed to elucidate the role of Spartan in the repair of radical-induced DPCs in the tissues of wild type and *SPRTN* hypomorphic (*SPRTN* $^{H/H}$) mice.[429]

4.2. Materials and methods

Chemicals and Reagents

2'-deoxythymidine, Boc-L-Tyr-OH, trifluoroacetic acid, and Nuclease P_1 from *Penicillium citrinum* were purchased from Sigma-Aldrich (St. Louis, MO). Glacial acetic acid, hydrochloric acid, and acetonitrile were obtained from Fisher Scientific (Buchs, Switzerland). Cell Lysis Solution, Protein Precipitation Solution, and RNaseA were purchased from Qiagen (Hilden, Germany). Amicon 3K filters were purchased from Millapore (Darmstadt, Germany). Omega Nanosep 10K Omega filters were purchased from PALL Life Science (Port Washington, NY).

Animal safety statement

Spartan hypomorphic mice (*SPRTN*$^{H/H}$) were obtained as described previously.[429] Animals were housed in standard rodent cages with 12-hour light-dark cycles and were given food and water *ad libitum*. Age-matched *SPRTN*$^{H/H}$ and wild type mice were euthanized by carbon dioxide inhalation and the liver, spleen, kidney, heart, thymus, and brain were excised for future study. All experimental procedures were approved by the Mayo Clinic Institutional Animal care and Use Committee (IACUC). In addition, all studies were conducted in accordance with the United States Department of Health and Human Services National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Cytotoxicity of DPC-inducing agents in Spartan proficient and Spartan deficient cell lines (MEF5 and MEF7)

Wild type MEF (MEF5, *SPRTN* $^{+/+}$) and isogenic *SPRTN*-deficient (MEF7, *SPRTN* $^{f/-}$) cells were obtained and used with permission from Dr. Yuichi Machida's laboratory (Mayo Clinic, MN).[429] The MEF7 cell lines were originally obtained from E13.5 embryos produced by crossing hemizygous (*SPRTN* $^{+/-}$) and floxed (*SPRTN* $^{f/+}$) mice and immortalized by serial passage.[429] The inducible Cre-ER^{T2} element was incorporated by infection with retroviral vectors and selected with 3 μg/μL puromycin.

MEF5 and MEF7 cells were plated in 100 μL Dulbecco's Modified Eagle's Medium containing 10% FBS at a density of 1 x 10^4 cells/well. Cells were permitted to adhere in an incubator humidified atmosphere of 5% carbon dioxide, 95% air, at 37 °C for 24 hours. Cells were then treated with 0 – 100,000 μM hydrogen peroxide or 0 – 1000 μM

cisplatin, DEB, or PM in 100 μL serum-free medium for 3 h at 37 ºC. Following treatment, cell media was replaced with 200 μL fresh FBS-containing media, and the cells were incubated an additional 48 h at 37 ºC. Cell viability was determined using an Alamar Blue assay[441] using a Synergy HI Microplate reader (BioTek, Winooski, VT).

Cytotoxicity of hydrogen peroxide in MEF5 and MEF7 cells after pharmacological inhibition of proteasome

MEF5 and MEF7 cells were plated in 100 μL Dulbecco's Modified Eagle's Medium containing 10% FBS at a density of 1 x 10^4 cells/well. Cells were permitted to adhere in an incubator humidified atmosphere of 5% carbon dioxide, 95% air, at 37 °C for 24 hours. The media was replaced with 100 μL of 10 μM lactacystin or 1 μM MG-132 in DMEM media, and the cells were incubated for 1 hour. Concentrations of lactacystin and MG-132 were selected based on previously calculated IC_{50} values for 26S proteasome inhibition. Cells were then treated with 0 – 100,000 μM hydrogen peroxide in 100 μL serum-free medium for 3 h at 37 ºC. Following treatment, cell media was replaced with 200 μL fresh FBS-containing media, and the cells were incubated an additional 48 h at 37 ºC. Cell viability was determined using an Alamar Blue assay[441] using a Synergy HI Microplate reader (BioTek, Winooski, VT).

DPC formation in HeLa cells and mice subjected to ionizing radiation (IR) treatment

HeLa cells (in triplicate) were plated in 10 mL Dulbecco's Modified Eagle's Medium containing 10% FBS at a density of 1 x 10^6 cells per 10cm dish. Cells were permitted to adhere in an incubator humidified atmosphere of 5% carbon dioxide, 95% air

until 50% confluency was reached, followed by treatment with 8 gy of ionizing radiation on a Precision X-Ray X-RAD 320 biological irradiator (North Branford, CT) for 174 seconds. Cells were immediately irradiated with 8 gy for an additional 174 seconds for 16 gy exposure. After irradiation, cells were collected and IR-induced DPCs were extracted as described below. As a control, HeLa cells (in triplicate) were cultured and left untreated before harvesting.

To investigate IR-induced DPC formation *in vivo*, C57BL/6 wild type mice were irradiated in a ^{137}Cs Mark I-68 irradiator (JL Shepherd & Associates, San Francisco, CA) at a rate of 137.44 cGy/min for 8.73 min to yield a 12 gy dose. After irradiation, the animals were euthanized and the livers and brains were immediately frozen at -20 ºC. IR-induced DPCs were extracted from the tissues as described below.

Cell extraction from wild type and *SPRTN*$^{H/H}$ mouse tissues (Scheme 1)

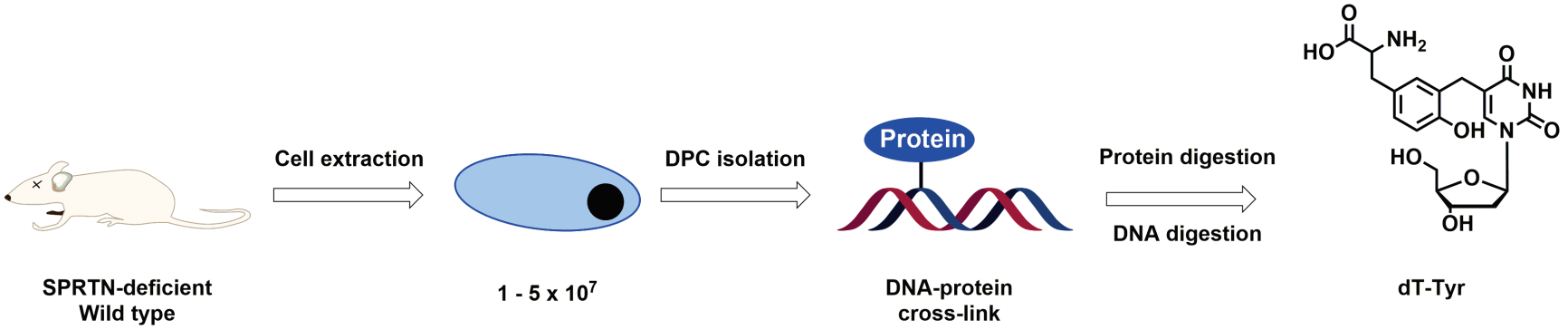
Brain, heart, liver, thymus, spleen, and kidney tissues from 10-month old wild type and hypomorphic (*SPRTN*$^{H/H}$) mice (100 – 800 mg, 3 per group) were diced using a razor blade and homogenized in 2 mL of sucrose buffer (20 mM phosphate buffer, pH 7.2, 250 mM sucrose) using gentle pipetting. The resulting homogenate was incubated on ice for 10 minutes, followed by centrifugation at 1500 RPM at 4 ºC for 10 min. The pelleted cells were re-suspended in 2 mL cell lysis solution (Qiagen, Hilden, Germany) and DNA/DPCs were extracted as described below.

DPC extraction from cells extracted from mouse tissues and IR-treated HeLa cells (Scheme 4-1)

IR-exposed HeLa cells (~ 1.0 x 10^7) or cells extracted from mouse tissue (~ 1.0 x 10^7) were re-suspended in 2 mL of cell lysis solution (Qiagen, Hilden, Germany) and pipetted until the cells were completely lysed. For cells obtained from brain and liver tissue (~ 1.0 x 10^8), the cells were typically re-suspended in 3 mL of cell lysis solution and processed as described above. The resulting lysates were treated with 40 μg RNaseA (10 μL, Qiagen) overnight at room temperature with gentle inversion. Following incubation, each sample was treated with 8 Units proteinase K (10 μL, Qiagen) and incubated overnight at room temperature with gentle inversion. Following incubation, each solution was treated with 750 μL of protein precipitation solution (Qiagen, Hilden, Germany) and vortexed for 1 min to yield a white turbid solution. Each sample was then centrifuged at 1500 RPM for 10 min to pellet the free proteins.

The supernatant containing the DNA/DPCs was decanted, and an equal volume of cold 100% EtOH was added to precipitate DNA/DPCs. Precipitated DNA/DPCs were washed with 70% EtOH twice and 100% EtOH, followed by re-suspending in 10 mM Tris buffer, pH 7.5. Successful extraction of DNA was confirmed by UV spectrophotometry at $\lambda = 260$ nm, and DNA concentrations were determined by dG analysis described in **Section 3.2**.

Scheme 4-1: Methodology for extracting ROS-induced or IR-induced DNA-protein cross-links. Tissues from either wild type (*SPRTN*$^{+/+}$) or hypomorphic (*SPRTN*$^{H/H}$) mice are homogenized to pellet cells. The recovered cells, or cells obtained from *in-vitro* experiments, are then lysed and DPCs are extracted following a previously published procedure. Isolated DPCs can then be enzymatically digested to a single nucleoside and amino acid (dT-Tyr), an adduct that can be quantified by our developed isotope dilution tandem mass spectrometry assay.



Quantitation of dT-Tyr obtained from wild type or SPRTN$^{H/H}$ tissues or IR-treated MEF cells

DPC-containing DNA extracted from the MEF cells and animal tissues as described above was re-suspended in 100 μL 10 mM Tris, pH 7.5 and treated with proteases and nucleases to digest the protein and DNA constituents of DPCs as described in detail in **Section 3.2**. The total digestion product dT-Tyr were enriched by offline HPLC purification, and radical-induced DPC formation was quantified using the developed dT-Tyr nanoLC-ESI$^+$-MS/MS assay (**Section 3.2**).

4.3. Results

4.3.1. Sensitivity of Spartan deficient (MEF5) and Spartan proficient (MEF7) cells towards DPC inducing agents

MEF5 and MEF7 cells were genetically engineered as described previously.[429] MEF7 only contain one floxed copy of the SPRTN gene, and expresses less than 50% of MEF5 wild type Spartan levels.[429] Floxing of the *SPRTN* gene was confirmed by PCR amplification of the 527 base pair floxed allele, and decreased Spartan expression was confirmed by western blot quantitation as described previously.[429]

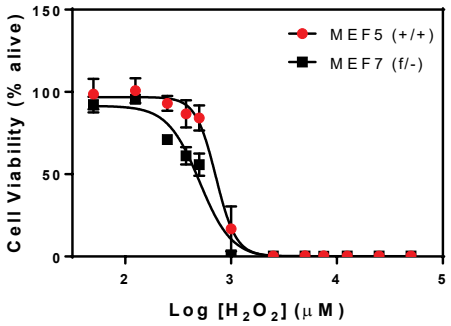
To compare the sensitivity of Spartan deficient (MEF5) and Spartan proficient (MEF7) towards known DPC inducing agents, cells (in quadruplet) were treated with increasing amounts of hydrogen peroxide, 1,2,3,4-diepoxybutane, cisplatin, phosphoramide mustard, and DEB from 0 to 2000 μM (**Figure 4-1A – D**). These agents were selected because they had been previously shown to form covalent DPCs in cells.[288, 289, 399, 525] Final concentration ranges were determined based on previous cell viability

results with other mammalian cells[288, 289, 399] and preliminary investigations. Cell viability and IC_{50} values were determined using the commercially available Alamar Blue assay.[441]

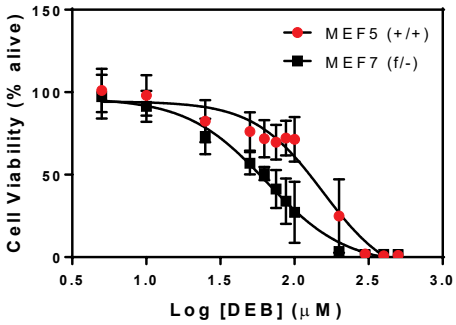
Overall, cells expressing lower amounts of Spartan (MEF7) were more sensitive towards DPC forming agents than their Spartan proficient counterparts (**Table 4-1**). When treated with hydrogen peroxide, MEF7 cells were 1.5-fold more sensitive than MEF5 cells (IC_{50} = 561.8 ± 69 and 839.3 ± 28.4 respectively. Similarly, MEF7 cells were also 1.5-fold more sensitive to PM (IC_{50} = 630.0 ± 46) compared to MEF5 cells (IC_{50} = 963.9 ± 62.1) and 1.3-fold more sensitive to cisplatin (IC_{50} = 39.1 ± 5.6) compared to MEF5 cells (IC_{50} = 50.4 ± 4.3). Finally, MEF7 cells were 2.4-fold more sensitive to DEB (IC_{50} = 64.7 ± 12.7) compared to MEF5 cells (IC_{50} = 155.4 ± 49.8).

Figure 4-1: Cytotoxicity of known DPC-inducing agents H_2O_2 (**A**), DEB (**B**), PM (**C**), and cisplatin (**D**) in Spartan deficient (MEF5) and Spartan proficient cells (MEF7).

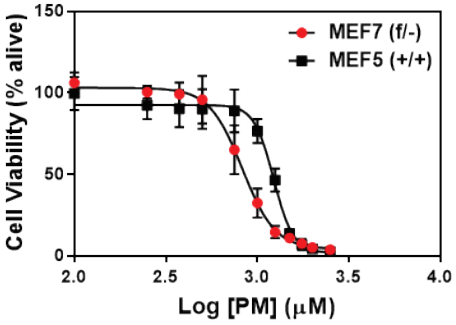
A.



B.



C.



D.

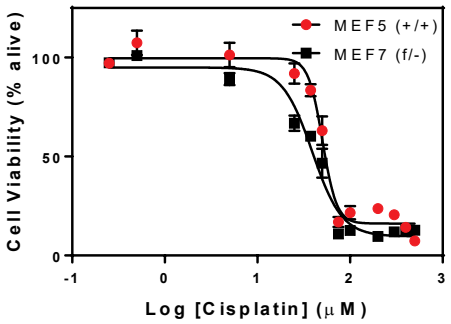


Table 4-1: Calculated IC_{50} values of MEF5 and MEF7 cells treated with 0 – 1000 μM cisplatin, PM, or DEB, or 0 – 100,000 μM hydrogen peroxide (H_2O_2).

Agent	MEF5 IC_{50} (μM)	MEF7 IC_{50} (μM)	Difference
H_2O_2	717.3 ± 59.9	504.0 ± 44	1.4 fold
Cisplatin	50.4 ± 4.3	39.1 ± 5.6	1.3 fold
Phosphoramide mustard	963.9 ± 62.1	630.0 ± 46	1.5 fold
DEB	155.4 ± 49.8	64.7 ± 12.7	2.4 fold

4.3.2. Effects of proteasome inhibitors on cell sensitivity towards DPC forming agents

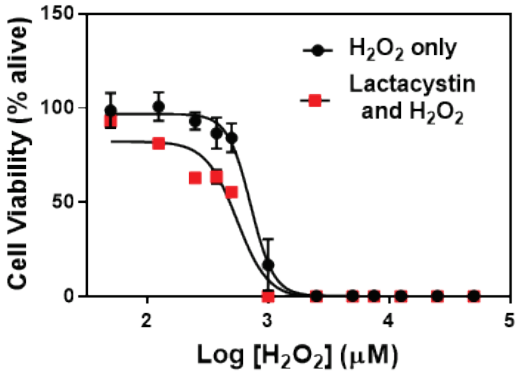
As discussed above, 26S proteasomal degradation has been previously reported to facilitate DPC repair by reducing the size of the lesion.[309, 427, 522, 526] However, others have failed to observe any effect of proteasomal inhibitors on DPC repair kinetics.[424] It is possible that the structural identity of DPC lesions affects their ability to be processed via proteasomal degradation. To determine the role of proteasome in repair of DPCs of varied structure, MEF5 and MEF7 cells were pretreated with proteasome inhibitor MG-132 (1 μM) or lactacystin (10 μM). MG-132 is a tripeptide aldehyde (Z-Leu-Leu-Leu-al) that inhibits catalytic subunits of the 26-proteasome complex,[527] and at higher concentrations (3 – 10 μM) retards the growth of mammalian cells by promoting cell cycle arrest.[528] Lactacystin is a selective and irreversible beta-lactone that inhibits the 20-S proteasome complex.[529, 530] Following treatment of MEF cells with hydrogen peroxide, cell viability was measured as described above to determine if the inhibition of the proteasome further sensitized cells to ROS-induced DPCs.

When MEF5 cells were pretreated with MG-132, there was no difference in cell cytotoxicity to hydrogen peroxide treatment only (IC_{50} = 717.3 ± 59.9 vs IC_{50} = 764.2 ± 224.4 **Figure 4-2A**). MEF7 cells pretreated with MG-132 showed a slight decrease in cell cytotoxicity to hydrogen peroxide treatment only (IC_{50} = 504.0 ± 44 and 469.3 ± 48.1 respectively, **Figure 4-2B**). Similarly, pretreatment with lactacystin modestly sensitized both cell types to hydrogen peroxide treatment. MEF5 IC_{50} values decreased from 717.3 ± 59.9 to 552.2 ± 65.7 (**Figure 4-3A**), while MEF7 IC_{50} decreased from 504.0 ± 44 to 475.6

± 21.7 respectively (**Figure 4-3B**). Taken together, inhibition of the proteasome does not sensitize both cell lines to hydrogen peroxide-induced DNA damage.

Figure 4-2: Cytotoxicity of hydrogen peroxide in MEF5 (**A**) and MEF7 cells (**B**) with and without pretreatment with proteasome inhibitor lactacystin (10 µM).

A.



B.

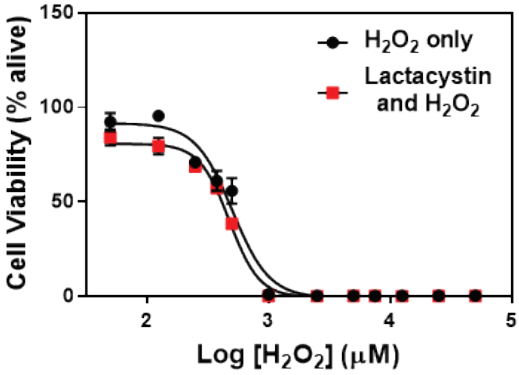
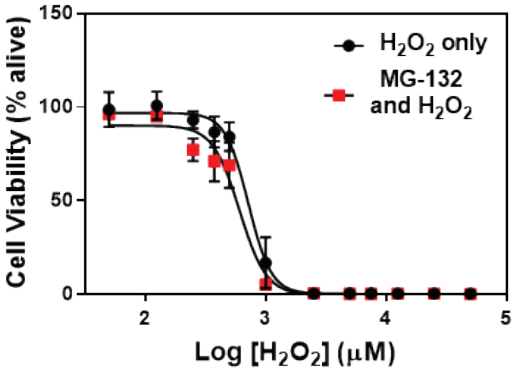
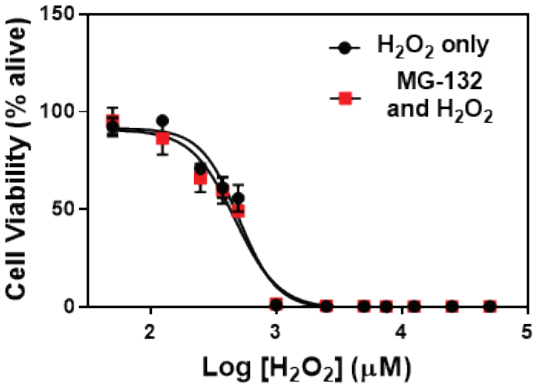


Figure 4-3: Cytotoxicity of hydrogen peroxide in MEF5 (**A**) and MEF7 (**B**) cells after pretreatment with proteasome inhibitor MG-132 (1 μM).

A.



B.



4.3.3. DPC formation in cells and tissues treated with ionizing radiation (IR)

To detect the formation of radical-induced DPCs in cells exposed to IR, HeLa cells were treated with 0, 8, and 16 gy of irradiation. Immediately following treatment, DPCs were extracted and enzymatically digested to release dT-Tyr adducts, which were quantified by isotope dilution HPLC-ESI-MS/MS using our previously validated methodology (Groehler *et al*, *Free Radic. Biol. Med.,* 2018). A linear increase of dT-Tyr concentrations was observed in IR treated cells, with 3.36 ± 0.79 and 6.32 ± 0.17 adducts per 10^8 nucleotides observed in cells treated with 8 and 16 gy IR, respectively (**Figure 4-4**). dT-Tyr levels in HeLa cells not exposed to IR were below the method's limit of quantitation (2 fmol).

We next investigated whether our nanoLC-ESI^+-MS/MS assay could be utilized to detect IR-induced DPC formation *in vivo.* C57BL/6 wild type mice were treated with 12 gy of IR, followed by extracting and processing IR-induced DPCs from the livers and brains as described above. We were able to detect dT-Tyr in both the brain (5.4 and 8.5 per 10^8 nucleotides) and liver (5.4 and 6.63 dT-Tyr per 10^8 nucleotides, **Figure 4-5**), confirming our assay could detect IR-induced DPC formation *in vivo*.

Figure 4-4: Quantitative analysis of radical-induced DPC formation in MEF cells exposed to 0, 8, and 16 gy of ionizing radiation.

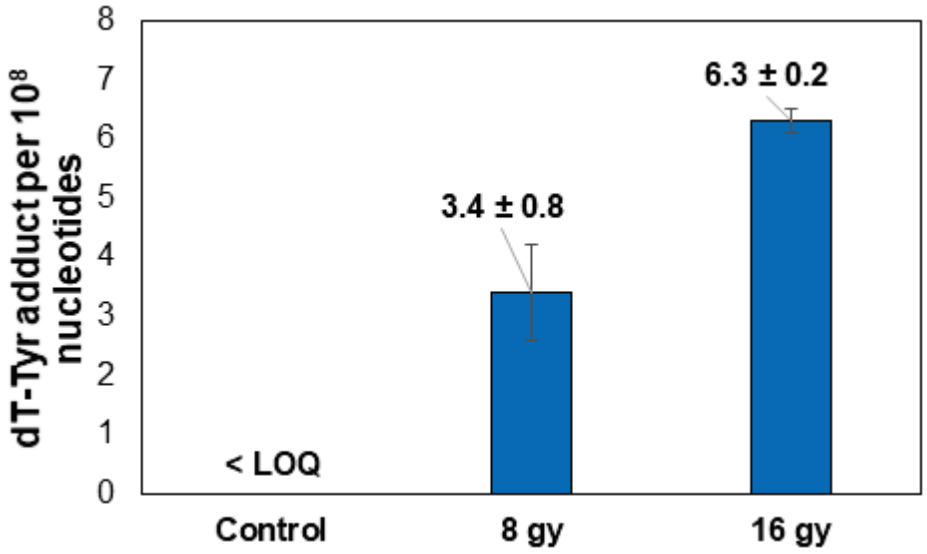
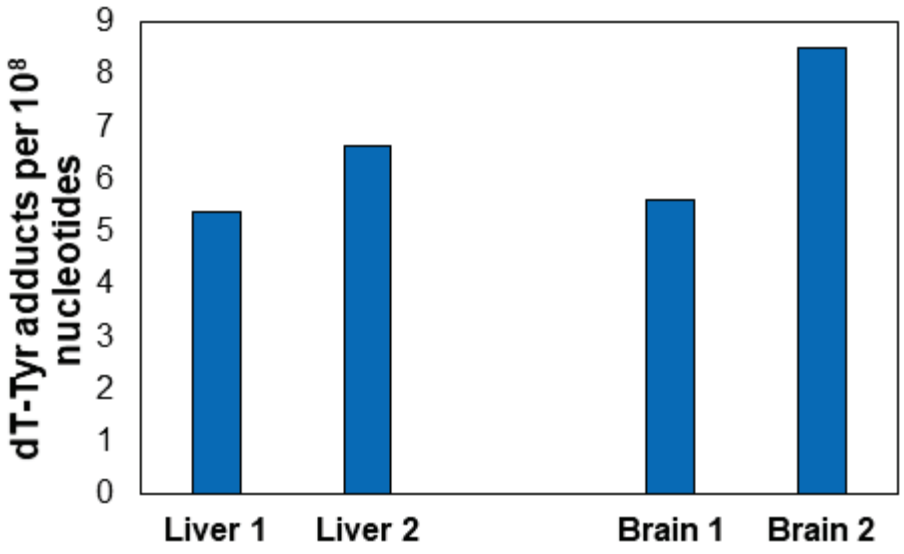


Figure 4-5: Quantitative analysis of radical-induced DPC formation in the brains and livers of mice exposed to 12 gy of ionizing radiation.



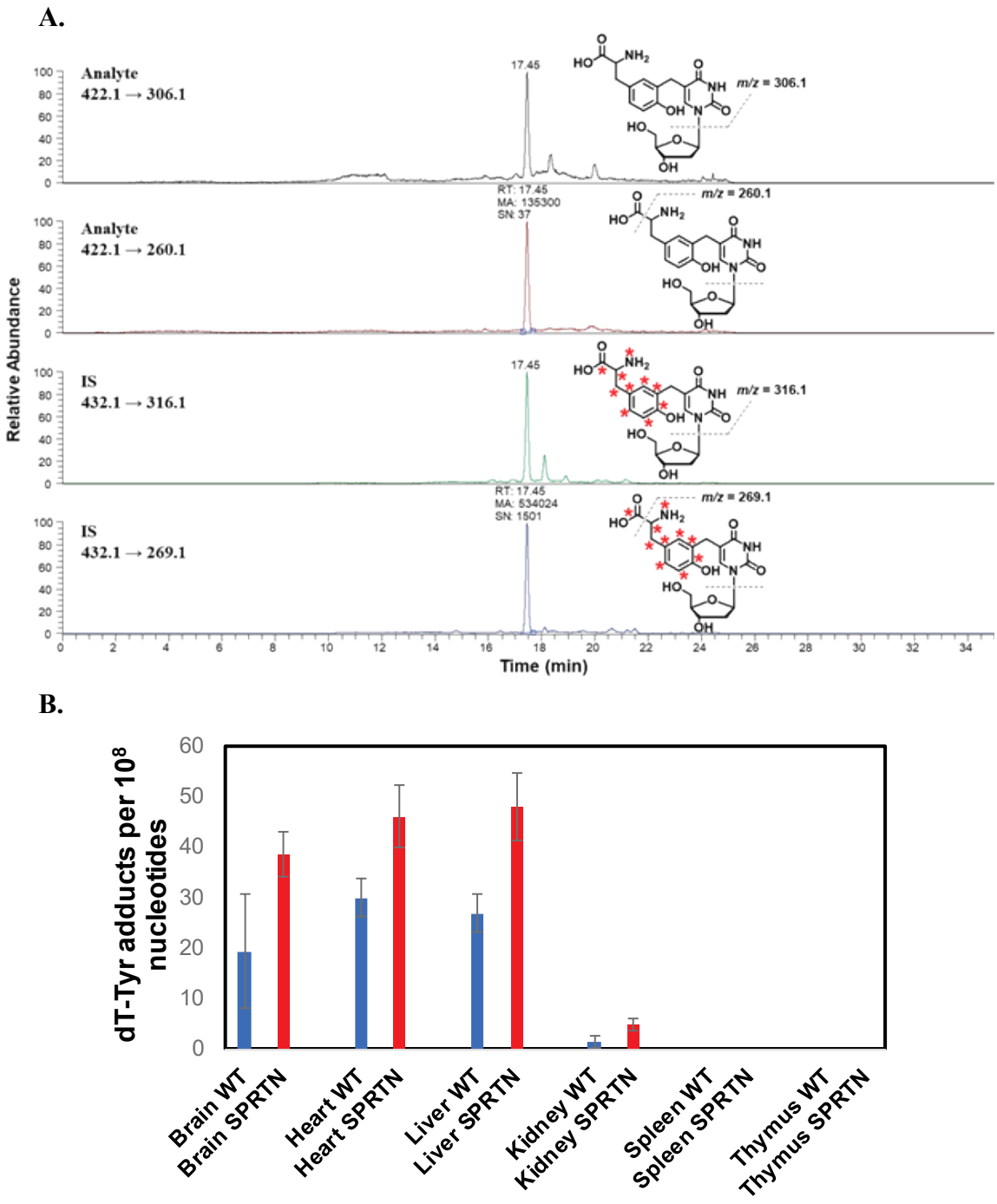
4.3.4. Quantitation of ROS-induced DPCs in the tissues of wild type and SPRTN$^{H/H}$ mice

To investigate the contribution of Spartan to repair endogenous DPCs, dT-Tyr conjugates were quantified in DNA isolated from liver, brain, thymus, spleen, kidney, and heart tissues of 10-month old wild type and SPRTN$^{H/H}$ (hypomorphic) mice (**Figure 4-6A**). Hypomorhic SPRTN$^{H/H}$ mice express only 10% of Spartan compared to wild type.[429] Previous studies have shown that Spartan-deficient mice experience rapid ageing, genome instability, kyphosis, and early development of cataracts.[429] Furthermore, recent study has revealed that SPRTN$^{H/H}$ mice accumulate significantly higher levels of Topoisomerase-DNA cross-links over time compared to wild type mice.[514] nanoLC-ESI-MS/MS detected significant amounts of dT-Tyr analyte (1.39 ± 1.05 to 47.91 ± 6.64 dT-Tyr per 10^8 nucleotides) in the liver, brain, heart, and kidneys of both strains of mice, while no detectable levels of dT-Tyr analyte were observed in the spleen and thymus (**Figure 4-6B**).

Quantitative analysis of the liver, brain, heart, and kidney DNA revealed that hypomorhic SPRTN$^{H/H}$ mice had significantly higher levels of dT-Tyr analyte compared to the wild type mice (**Figure 4-6B**). The highest levels of dT-Tyr were observed in the liver, with 47.91 ± 6.64 and 26.84 ± 3.72 adducts per 10^8 nucleotides in SPRTN$^{H/H}$ and wild type mice respectively, and in the heart, with 46.02 ± 6.08 and 29.89 ± 3.68 adducts per 10^8 nucleotides in SPRTN$^{H/H}$ and wild type mice respectively (**Figure 4-6B**). Analysis of the brain revealed 38.48 ± 4.37 and 19.34 ± 11.37 dT-Tyr adducts per 10^8 nucleotides in SPRTN$^{H/H}$ and wild type mice, respectively (**Figure 4-6B**). The greatest difference in dT-Tyr adduct levels were found in the kidney, with 4.81 ± 1.12 and 1.39 ± 1.05 adducts per 10^8 nucleotides in SPRTN$^{H/H}$ and wild type mice, respectively (**Figure 4-6B**). Overall,

these results indicate that while Spartan contributes to DPC repair in mouse tissues, another proteolytic pathway must be also in place to allow for DPC removal in the absence of Spartan.

Figure 4-6: Representative trace of dT-Tyr adduct detected in the heart of SPRTN-deficient mouse (**A**). Quantitation of dT-Tyr adducts in the brain, heart, liver, kidney, thymus, and spleen of 10-month-old wild type (*SPRTN*$^{+/+}$, blue bars) and hypomorphic (*SPRTN*$^{H/H}$, red bars) mice (**B**).



4.4. Discussion

Ionizing radiation is a high energy radiation (alpha particles, beta particles, and x-rays) that induces the release of electrons from molecules, which generates ions that can break covalent bonds.[523, 524] IR exposure can directly damage DNA by creating ions that physically break the phosphodiester backbone, yielding lesions such as double strand breaks.[523, 524] IR exposure can also damage DNA indirectly through the formation of hydroxyl radicals,[523, 524] which can then oxidatively damage positions on DNA nucleosides and residues of proteins. Therefore, IR exposure could potentially lead to hydrogen abstraction from the methyl position of a thymidine nucleoside and subsequently form a cross-link with a neighboring tyrosine residue as observed with hydroxyl radicals.[362, 363, 481]

Interestingly, the same types of thymine-tyrosine DNA-protein cross-links can also form endogenously in unexposed cells and tissues. ROS-induced DPCs are formed after hydroxyl radicals abstract a hydrogen from the 5-methyl position of thymidine to yield a reactive thymidine radical, which undergoes a one-electron addition to the 3-position of a tyrosine to yield a stable methylene linkage (**Scheme 3-2**).[481, 482] Subsequent hydrogen abstraction from the 3-position of tyrosine re-aromatizes the phenol ring to yield the stable DNA-protein cross-link (**Scheme 3-2**).[481, 482]

We recently developed a sensitive nanoLC-ESI$^+$-MS/MS assay for the total digestion product, 2-amino-3-(4-hydroxy-3-((1-((2R,4R,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydro-furan-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)methyl)phenyl)propanoic acid (dT-Tyr), to perform absolute quantitation of hydroxyl radical-induced DPC formation in cardiomyocytes, (Groehler *et al*, *Free Radic. Biol. Med.,*

2018). Using our developed dT-Tyr nanoLC-ESI$^+$-MS/MS assay, we reported that radical-induced DPC formation increased in the cardiomyocytes of rats that underwent LAD ligation/reperfusion as compared to sham surgery (35.7 ± 2.2 vs 18.2 ± 1.5 dT-Tyr adducts per 10^8 nucleotides respectively, **Figure 3-2B**). We concluded that ischemia-reperfusion injury increased the levels of hydroxyl radicals in affected myocardium, leading to an increase in radical-mediated DPCs in cardiomyocyte DNA as described in **Section 3.3.2**.

In the present study, we investigated the respective roles of Spartan and the 26S proteasome in the repair of ROS and IR-induced DPCs. We first confirmed that our previously developed dT-Tyr MS/MS assay, which has successfully been used to quantify ROS-induced DPC formation in cardiomyocytes, could be used to quantify IR-induced DPC formation. After treating HeLa cells with 0, 8, and 16 gy of IR, we observed a linear increase (0 – 6.32 ± 0.17) in dT-Tyr per 10^8 nucleotides (**Figure 4-4**). Furthermore, we were able to quantify dT-Tyr in the livers and brains of mice treated with 12 gy of IR (**Figure 4-5**). Together, these experiments confirmed our assay could be used to quantify IR-induced DPC formation in both cells and tissues, and we could investigate DPC repair of both ROS and IR-induced DPCs.

To investigate the role of Spartan in protecting cells against the cytotoxic effects of DPCs, cell cytotoxicity experiments were conducted with MEF5 (wild type) and MEF7 (*SPRTN* $^{f/-}$) cells. These experiments revealed that the Spartan-deficient cells were modestly more sensitive to the DPC-inducing agents hydrogen peroxide, cisplatin, phosphoramide mustard (PM), and 1,2,3,4-diepoxybutane (DEB) (**Table 4-1**). These results support a role for Spartan in removing toxic DPCs induced by cross-linking agents. However, it should be noted that all four of the agents investigated here are known to also

induce significant levels of nucleoside monoadducts,[31, 34, 201, 210, 327, 383-385, 531-533] as well as DNA intrastrand cross-links[215, 534-536] and interstrand cross-links, which can contribute to the observed toxicity.[215-218, 327, 384, 537, 538] For example, Kim *et al* discovered that the depletion of Spartan by siRNA leads to an increase in error prone bypass across the UV-induced thymine dimers.[539, 540] Therefore our toxicity experiments must be interpreted with caution.

Inhibitors of the 26S proteasome (MG-132 or lactacystin) had minimal effects on toxicity of hydrogen peroxide in MEF5 and MEF7 cells (**Figure 4-3** and **4-4**). These results are consistent with previous results by Nakano *et al,*[424, 467] who failed to observe ubiquitination of FA-induced DPCs, which is required for proteolytic digestion by the 26S proteasome. In contrast, Baker *et al* and Quievryn *et al* observed a reduced repair of a Dmnt1-containing plasmid[427] and FA- and Cr(IV)-induced DPCs[309, 522] in cells when the 26S proteasome was pharmacologically inhibited. Furthermore, Maskey *et al* showed that a deficiency in Spartan also lead to an increase in Top1ccs, indicating that the 26S proteasome cannot independently digest topoisomerases ccs to allow the excision of this proteasome work together in the repair of this DPC.[514] These discrepancies could be due to the different DPC structures examined.

To directly test the role of Spartan in repair of endogenous ROS induced DPCs *in vivo*, we employed wild type (*SPRTN*$^{+/+}$) and hypomorphic (*SPRTN*$^{H/H}$) mice. nanoLC-ESI-MS/MS analysis of dT-Tyr conjugates in DNA isolated from brain, liver, spleen, heart, thymus, and kidney tissues of the two strains of mice revealed a 1.5 to 2-fold increase in dT-Tyr adduct in the liver, brain, kidney, and heart of hypomorphic (*SPRTN*$^{H/H}$) mice compared to wild type (*SPRTN*$^{+/+}$) mice (**Figure 4-6**). Previous investigations by Vaz *et al*

observed a 3-fold increase in endogenous DPCs in Spartan-depleted HeLa cells (CRISPR-Cas9-created *SPRTN* partial knockouts) compared to wild type cells using silver staining technique.[513] When endogenous DPC levels in the same HeLa cells were quantified by the SDS/K^+ precipitation assay, a similar 1.5-fold increase was observed in Spartan-depleted cells compared to the wild type.[513] Taken together with our dT-Tyr results, these results suggest that Spartan is directly involved in repair of radical-induced DPCs, although additional proteases may be involved in DPC repair.

Previous investigations have detected ROS-mediated DPCs in unexposed animals. For example, Izzotti *et al* observed an age-related increase in ROS-induced DPC formation and the ROS biomarker 8-oxoguanine in the liver, brain, and heart of mice.[270] Similarly, Leutner *et al* observed an age-related increase in the lipid peroxidation biomarker malondialdehyde coupled with an increase in activity of CuZn-superoxide dismustase and glutathione reductase in the brain,[541] providing additional evidence that hydroxyl radical concentration increases in brain tissues over time. Finally, the filtration of blood by kidneys exposes glomeruli cells to particulate and soluble stimuli, such as phorbal 12-myristate 13-acetate (PMA)[542] and arachidonic acid,[543] which induces phagocytotic cells to release peroxides and hydroxyl radicals.[544] However, to our knowledge, our study in the first to accurately quantify DPC formation in unexposed mice.

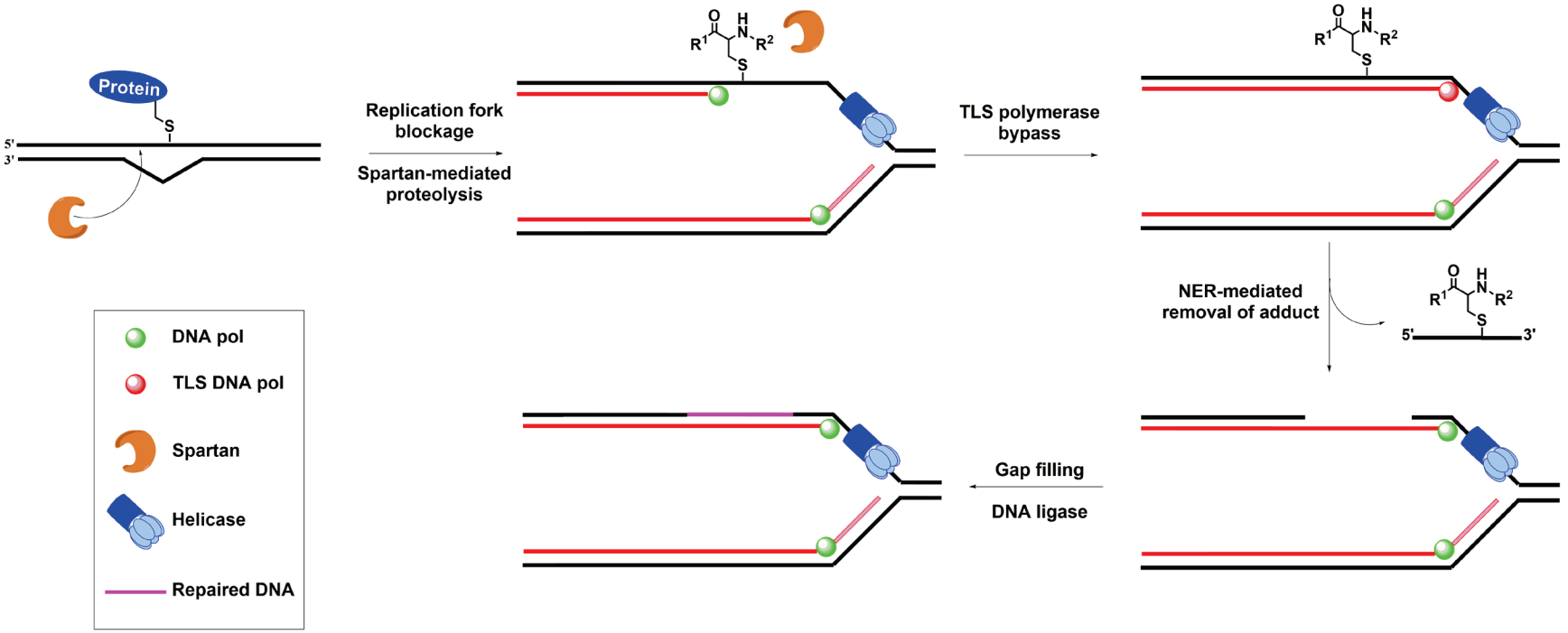
Our failure to detect dT-Tyr in DNA isolated from mouse spleen and thymus tissues (**Figure 4-6**) suggests that the specialized lymphatic immunoproteasomes (i-proteasomes) may be involved in repair of radical-induced DPCs. I-proteasomes are defined by tissue-specific inducible catalytic subunits of the low-molecular-weight protein 2 (LMP2), LMP7, and multicatalytic endopeptidase complex subunit-1 (MECL-1).[545] The primary function

of i-proteasomes is to digest proteins to peptides with a hydrophobic C terminus, which can be processed to recognize major histocompatibility (MHC) class 1 molecules.[545-547] Displaying these processed peptides on the membranes of CD8 T-cells allows for the detection of bacterial and viral pathogens.[545-547] Recently, additional functions of i-proteasomes, including regulating cellular homeostasis,[545, 548] tumor development,[549] acute retinal injury,[550-552] and insulin-dependent diabetes mellitus/diabetes insipidus,[553] have been elucidated in both humans and mice. Most interestingly, Ding *et al* observed an increase in oxidatively damaged proteins in the tissues of i-proteasome knockout (LMP2 KO) mice.[554] Similarly, Ferrington *et al* observed that retinal pigment epithelial cells from LMP7/MECL-1 double knockout mice were significantly more sensitive to oxidative stress in culture.[550] Taken together, i-proteasomes expressed in lymphatic tissues can be transported to other tissues to protect them from oxidative damage and digest any oxidatively damaged proteins, potentially including the radical-induced DPCs investigated in this manuscript. We are currently investigating whether i-proteasomes are directly involved in the repair of ROS-induced DPCs by quantifying dT-Tyr adducts in the tissues of LMP2 and LMP2/LMP7 KO mice.

One proposed model of DPC repair in mammalian systems couples the recruitment of proteases to a DPC with stalled DNA replication.[428, 431, 512] In this model, the replisome/CMG helicase becomes stalled after encountering a DPC,[264, 265, 421] which leads to the recruitment of Spartan (or any protease) to the lesion.[427, 430, 431, 514] Once recruited, Spartan will bind to ssDNA regions created by the bulky nature of DPCs, bringing Spartan into proximity of its substrate and allowing for proteolytic digestion of the protein constituent (**Scheme 4-2**).[431] The resulting DNA-peptide lesion can then be immediately

bypassed by TLS polymerases,[430, 431] and eventually repaired by NER (**Scheme 4-2**).[309, 426, 427, 555] Recently, Chesner *et al* utilized strand-specific primer extension-quantitative polymerase chain reaction (SSPE-qPCR) to show that the repair of a plasmid containing human oxoguanine glycosylase 1 (OGG1) occurred within 2 hours post-transfection in wild type (V79) and NER-deficient (VH1) CHO cells, well before DNA replication would become a factor.[555] Combined with our results suggesting that multiple proteases are required for the digestion of the protein constituent, it is unlikely the stalling of the replisome is the only molecular signal to stimulate DPC repair.

Scheme 4-2: Proposed Spartan-induced protein digestion mechanism for DPC removal. After the DPC is recognized, the metalloprotease Spartan is recruited to the area and binds a ssDNA region induced by the bulkiness of the lesion. The resulting DNA-peptide cross-link can then be immediately bypassed by TLS polymerases if DNA replication is occurring, of the smaller adduct can be excised via NER.



5. EB-FAPy-dG Adducts of 1,3-Butadiene: Synthesis, Structural Identification, and Detection in Human Cells

5.1. Introduction

1,3-butadiene (BD) is a recognized human and animal carcinogen present in cigarette smoke,[556] automobile exhaust,[177] and wood burning fumes.[557] BD is also an important industrial chemical broadly used in polymer and rubber production, with 519,400 workers exposed to BD in the United States (Bureau of labor statistics).[184] BD is metabolically activated by cytochrome P450 monooxygenases to three inherently reactive epoxides: 3,4-epoxy-1-butene (EB), 1,2,3,4-diepoxybutane (DEB), and 1,2-dihydroxy-3,4-epoxybutane (EBD) (**Scheme 1-6**).[192] These electrophilic species can enter the cell nuclei and react with nucleophilic bases of DNA to form covalent nucleobase adducts.[558] BD-DNA adducts are thought to be responsible for its carcinogenicity, as they can lead to DNA polymerase errors during replication.[196, 197, 559-562] A number of BD-DNA adducts have been identified, including N-7-(2-hydroxy-3-buten-1-yl)-guanine (EB-Gua I),[200, 563] N-7-(1-hydroxy-3-buten-2-yl) guanine (EB-Gua II),[200, 563] N7-(2,3,4-trihydroxybut-1-yl)-guanine (N7-THBG),[206, 207] N1-(2,3,4-trihydroxybut-1-yl)-adenine (N1-THB-Ade),[564] 1,4-*bis*-(guan-7-yl)-2,3-butanediol (*bis*-N7G-BD)[565,566] and 1,N^6-(1-hydroxymethyl-2-hydroxypropan-1,3-diyl)-2′-deoxyadenosine (1,N^6-HMHP-dA) (**Figure 1-6**).[209, 567] Among these, N7-guanine adducts (EB-Gua and N7-THBG) are by far the most abundant

and have been detected in BD-treated laboratory animals,[202, 206, 565] occupationally exposed workers,[207] and current smokers.[207]

Despite their predominant formation, N7-guanine adducts of BD may not be the most biologically relevant because they retain the ability to pair with cytosine,[252, 568] and are hydrolytically labile because of the intrinsic destabilization of the glycosydic bond when the N7 position of guanine is alkylated.[33] Spontaneous depurination of N7-THBG and EB-Gua releases free base adducts form the DNA backbone and leads to the formation of abasic sites (**Scheme 1-8**). Alternatively, the C8 position of EB-Gua can be attacked by hydroxyl anions (**17, Scheme 1-8**), leading to imidazole ring opening and the formation of N^6-(2-Deoxy-D-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-N-(2-hydroxybut-3-en-1-yl)formamidopyrimidine (EB-Fapy-dG) adducts (**18, Scheme 1-8**).[233]

This imidazole ring scission dramatically changes the molecular shape and the hydrogen bonding characteristics of the parent nucleobase, modifying its base pairing preferences and leading to mispairing during DNA replication.[227, 228] Alternatively, FAPy adducts may be recognized by DNA repair enzymes and excised from the DNA backbone.[254, 256, 569-572] Unsubstituted (FAPy-dG) and alkyl-FAPy adducts (N^5-alkyl-FAPy-dG) are known substrates for the base excision repair (BER) enzymes 8-oxoguanine glycosylase (OGG1),[570, 571] endonuclease III-like protein (NTHL1),[570, 572] and endonuclease 8-like protein I – III (NEIL1, NEIL2, and NEIL3 respectively).[570, 572] Recently, NEIL1 has been shown to excise the bulky aflatoxin B_1-FAPy-dG adduct,[573] making it the primary candidate to recognize and excise any N^5-alkyl-FAPy-dG formed after exposure to the epoxide metabolites of BD.

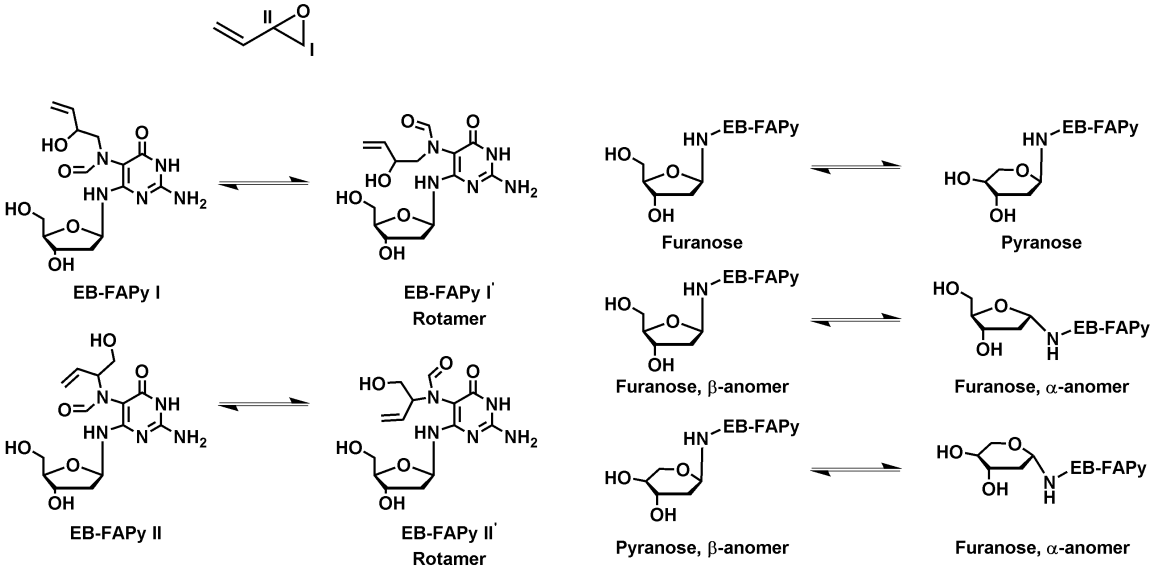
The most studied example of the N^5-alkyl-FAPy-dG adducts are those formed upon exposure to aflatoxin B_1 (AFB_1), a potent liver carcinogen produced by the common soil fungus *Aspergillus flavus.*[574, 575] Dietary exposure to AFB1 is recognized as a major risk factor for hepatocellular carcinoma in developing countries.[241, 576, 577] Metabolic activation of AFB_1 yields electrophilic AFB_1-*exo*-8,9-epoxide,[578] which alkylates DNA to form *trans*-8,9-dihydro-8-(N7-guanyl)-9-hydroxyaflatoxin B_1 (AFB_1-N7-dG), and the corresponding ring open AFB_1-β-FAPy adducts (Scheme 1B).[579] AFB_1-β-FAPy-dG adducts have been shown to induce high levels of G→T transversion mutations in *Escherichia coli*, nearly 6 times more frequent than G→T transversions caused by the cationic AFB_1-N7-dG adduct.[235] AFB1-FAPy-dG ability to induce G → T transversion mutations is consistent with the observed G→ T mutations in codon 249 of the p53 tumor suppressor gene observed in AFB1-treated human hepatocytes cultures.[249, 250] In addition, AFB1-induced G → T mutations in the *ras* oncogene contribute to liver tumor progression.[247]

In addition to AFB_1-β-FAPy, other N^5-alkyl-FAPy-dG reported *in vivo* include Me-FAPy-dG and Me-FAPy-dA induced by methylating agents and FAPy adducts of antitumor nitrogen mustards.[227, 233, 253, 580] When Me-FAPy lesions incorporated into M13mp18 vector were introduced in *E. Coli* bacteria, they blocked DNA replication and increased mutational frequency following SOS induction.[227, 253] Gruppi *et al* treated human mammary tumor (MDA-MB-321) cells with *bis*-(2-chloroethyl)ethylamine (NM) and structurally characterized two novel NM-FAPy lesions, NM-FAPy-dG adduct and FAPy-dG-NM-dG cross-link, which were measured at 180 and 6 adducts per 10^6 nucleotides respectively. Overall, previous studies indicate that N^5-alkyl-FAPy-dG adducts can form

in mammalian cells, however, their prevalence and biological significance is not fully understood.

We hypothesized that BD-induced EB-Gua I and EB-Gua II can form the corresponding EB-Fapy-dG adducts in the presence of hydroxyl anions (**Figure 5-1**). In the present study, authentic standards of 3,4-epoxy-1-butene (EB)-derived guanine Fapy adducts (EB-Fapy) were prepared for the first time, structurally characterized by NMR and mass spectrometry, and used to develop an isotope dilution mass spectrometry based methods for their quantitation in EB-treated DNA and human cell culture. pH and concentration dependence for adduct formation was investigated to determine whether EB-FAPy-dG could form under physiological conditions, potentially contributing to genotoxicity and mutagenic effects of BD.

Figure 5-1: Chemical complexity of EB-Fapy-dG adduct, showing all possible isomers, rotamers, and atropisomers.



5.2. Materials and methods

Note: *3,4-epoxy-1-butene (EB) is a known carcinogen and must be handled with adequate safety precautions.*

Chemicals and Reagents:

LC-MS grade water, methanol and acetonitrile were purchased from Fisher Scientific (Pittsburgh, PA). 3,4-epoxy-1-butene (EB), *N,N*-dimethylformamide dimethyl acetal, dimethoxytrityl chloride, trifluoroethanol, dichloroacetic acid, and calf thymus DNA were obtained from Sigma-Aldrich (Millwaukee, WI/St. Louis, MO). 2'-deoxyguanosine was obtained from Carbosynth (San Diego, Ca). Hypersep Hypercarb SPE cartridges (100 mg/3 mL) were obtained from Thermo Scientific (Waltham, MA). Omega Nanosep 10K filters were obtained from PALL Life Science (Port Washington, NY). All other chemicals and solvents were obtained from Sigma-Aldrich (Millwaukee, WI; St. Louis, MO) unless specified otherwise. Isotopically labeled $^{15}N_3$-dG was provided by Dr. Rajat Subhra Das and Professor Ashis Basu (University of Connecticut).

Cell culture of wild type mouse embryonic fibroblast (MEF) cells and isogenic knockouts of Endonuclease 8-like 1 (NEIL1)

Wild type MEF ($NEIL1^{+/+}$) and isogenic $NEIL1^{-/-}$ MEF cells were a gift from Dr. Robert Turesky's laboratory (University of Minnesota), used with permission from Dr. R. Stephen Lloyd (Oregon Health and Science University).[581] Both cell lines were seeded at a density of 1 x 10^6 cells in DMEM supplemented with 10% FBS. Cells were cultured in an incubator humidified atmosphere of 5% carbon dioxide, 95% air at 37 ºC.

Synthesis of N^6-(2-Deoxy-D-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-N-(2-hydroxybut-3-en-1-yl)formamidopyrimidine (EB-FAPy-dG, Scheme 5-1)

(E)-N'-(9-((2R,4S,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-6-oxo-6,9-dihydro-1H-purin-2-yl)-N,N-dimethylformimidamide (2'-dG-N^2-N,N-dimethylformimidamide) (**22**)

2'-Deoxyguanosine (1 g, 3.70 mmol) was dissolved in dry methanol (9.5 mL) under argon, followed by the dropwise addition of *N,N*-dimethylformamide dimethyl acetal (2 mL, 14.8 mmol) over 5 minutes. The resulting suspension was stirred at room temperature for 60 h. The precipitate was collected by vacuum filtration, and the solid was washed with cold methanol to yield a white solid (980 mg, 81.7% yield). ESI^+-MS/MS (**22**): 323.0 $[M + H]^+ \rightarrow 207.1 [M - dR + H]^+$. ^{1}H-NMR (500 MHz, DMSO-d6): 11.32 (s, N1-H, 1H), 8.56 (s, C8-H, 1H), 8.04 (s, N=CH-N(Me)$_2$, 1H), 6.25 (t, *J* = 7 Hz, C'1-H, 1H), 5.30 (d, *J* = 3 Hz, C'3-OH, 1H), 4.94 (t, *J* = 5 Hz, C'5-OH, 1H), 4.42 – 4.34 (m, C'3-H, 1H), 3.89 – 3.79 (m, C'4-H, 1H), 3.59 – 3.51 (m, C'5-H, 2H), 3.16 (s, N-(CH$_3$)$_2$, 3H), 3.04 (s, N-(CH$_3$)$_2$, 3H), 2.64 – 2.64 (qt, *J* = 7 Hz, C'2-H, 1H), 2.28 – 2.19 (ddd, *J* = 14, 6, 3 Hz, C'2-H, 1H). ^{13}C-NMR (400 MHz, DMSO-d6): 158.46, 158.06, 157.74, 150.10, 137.12, 120.16, 88.20, 83.24, 71.39, 62.29, 41.12, 35.12.

(E)-N'-(9-((2R,4S,5R)-5-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-hydroxytetrahydrofuran-2-yl)-6-oxo-6,9-dihydro-1H-purin-2-yl)-N,N-dimethylformimidamide (**23**)

Compound **22** (958 mg, 2.97 mmol) was dissolved in 20 mL anhydrous pyridine, and 300 mg of 4,4-dimethoxytrityl chloride (DMTCl) was added. Three aliquots of 300 mg

DMTCl was added to the solution every 15 minutes (1.2 g total, 2.97 mmol), and the solution was allowed to stir at room temperature for 4 h. The solution was concentrated *in vacuo* to yield a brownish red oil, which was purified by silica column chromatography using methanol in dichloromethane to yield pure **23** as a white solid (1.6 g, 86%). ESI^+-MS/MS (**23**): 625.0 $[M + H]^+ \rightarrow$ 207.1 $[M - dR + H]^+$ and 303.1 $[DMT^+]$. ^{1}H-NMR (500 MHz, $CDCl_3$): 8.93 (s, N1-H, 1H), 8.59 (s, C8-H, 1H), 7.72 (s, N=CH-N$(Me)_2$, 1H), 7.42 (d, J = 8 Hz, DMT-H, 2H), 7.36 – 7.26 (m, DMT-H, 8H), 6.82 (d, J = 8, DMT-H, 4H), 6.40 (t, J = 6 Hz, C'1-H, 1H), 4.68 – 4.58 (m, C'4-H,1H), 4.18 – 4.11 (m, C'3-H, 1H), 3.80 (s, DMT-OCH_3, 6H), 3.47 – 3.42 (m, C'5-H, 1H), 3.34 – 3.29 (m, C'5-H, 1H), 3.14 (s, N-$(CH_3)_2$, 3H), 3.08 (s, N-$(CH_3)_2$, 3H), 2.76 (s, C'3-OH, 1H), 2.62 – 2.54 (q, J = 5 Hz, C'2-H, 1H), 2.58 – 2.49 (m, C'2-H, 1H).

N^6-(2-Deoxy-D-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-N-(2-hydroxybut-3-en-1-yl)formamidopyrimidine (**25**)

Protected dG (**23**) (1 g, 1.6 mmol) was dissolved in 8.17 mL trifluoroethanol, and 1.93 mL of 3,4-epoxybutene (24 mmol, 10 equiv.) was added dropwise over 15 min. The solution was allowed to stir for 24 h at room temperature and concentrated *in vacuo* to yield a white gum. Intermediate **24** was isolated by silica column purification (eluent methanol in dichloromethane with 1% triethylamine) to yield an off-white solid (150 mg, 13.5% yield). ESI^+-MS/MS (**24**): 695 $[M + H]^+ \rightarrow$ 277.2 $[M - dR + H]^+$, 303.1 $[DMT^+]$, and 207.1 $[M - dR - EB + H]^+$.

Compound **24** (150 mg, 0.216 mmol) was dissolved in 100 µL trifluoroethanol and 100 µL 1 M NaOH, and the solution was stirred for 16 h at room temperature. The solution

pH was neutralized to pH 7.0 using 5% HCl, and the solution was concentrated under vacuum to yield an off-white solid. The resulting solid was chilled in an ice bath (0 °C) for 15 min and dissolved in 250 μL of 1% dichloroacetic acid in dichloromethane. The solution was stirred for 10 min and immediately neutralized to pH 7.0 using triethylamine. The solution was concentrated *in vacuo* to yield a reddish yellow solid.

The resulting crude product was re-suspended in 1 mL of water and purified by semi preparative RP-HPLC using Zorbax XDB C18 column (250 mm × 10 mm, 5 μ, Agilent Technologies, Palo Alto, CA) with a gradient of water and acetonitrile at a flow rate of 1 mL/min. Solvent composition was changed from 0 to 3% B over the course of 5 min, 3 to 5% B over the course of 5 min, 5 to 10% B over the course of 10 min, 10 to 75% over the course of 5 min, 75 to 0% over the course of 3 min, and held at 0% B over 7 min. Broad peaks eluting at 9.5, 10.7, and 11.8 min corresponded to the EB-FAPy base (**25a, Scheme 5-1**), and broad peaks eluting at 13.7, 14.1, and 14.8 min corresponded to the desired EB-FAPy nucleoside (**25b, Scheme 5-1**). The collected fractions were concentrated *in vacuo* and characterized by ESI^+-MS/MS, ^{1}H-NMR, and HSQC. $^{15}N_3$-EB-Fapy adducts were prepared analogously starting with 15 mg of $^{15}N_3$–dG to be used as internal standards for mass spectrometry.

N-(2,4-diamino-6-oxo-1,6-dihydropyrimidin-5-yl)-N-(2-hydroxybut-3-en-1-yl)-formamido-pyrimidine (**25a**): ESI^+-MS/MS (**27a**): 240.1 $[M + H]^+ \rightarrow$ 222.2 $[M - H_2O + H]^+$ and 194.1 $[M - H_2O - CO + H]^+$. ^{1}H-NMR (600 MHz, DMSO-d6): 10.34 (s, N1-H, 1H), 8.10 (s, HC=O, 0.5 H), 7.79 (s, HC=O, 0.5 H), 6.47 (s, N^2H_2, 2H), 6.34 (s, C4-NH_2, 2H), 6.13 (s, OH, 1H), 5.85 – 5.75 (m, CH=CH_2, 1H), 5.30 – 5.18 (dd, J = 17, 6 Hz,

CH=CH**2,** 1H), 5.15 – 5.00 (dd, *J* = 12, 5 Hz, CH=CH**2**, 1H), 4.79 – 4.69 (q, *J* = 7 Hz, CH-OH, 1H), 3.55 – 3.47 (m, CH_2, 2H).

N^6-(2-Deoxy-D-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-N-(2-hydroxybut-3-en-1-yl)formamidopyrimidine (**25b**) ESI^+-MS/MS: 356.1 $[M + H]^+ \rightarrow$ 240.1 $[M - dR + H]^+$, 222.2 $[M - dR - H_2O + H]^+$ and 194.1 $[M - dR - H_2O - CO + H]^+$. ^{1}H-NMR (600 MHz, DMSO-d6): 10.34 (s, N1-H, 1H), 7.90 (s, HC=O, 0.25 H), 7.83 (s, HC=O, 0.25 H), 7.77 (s, HC=O, 0.5 H), 6.45 (s, N^2H_2, 2H), 6.35 (s, C4-NH, 1H), 6.0 (s, C'5-OH, 0.5 H) 5.85 – 5.75 (m, CH=CH_2, 1H), 5.80 (s, C'5-OH, 0.5 H), 5.75 (s, C'3-OH, 0.5 H), 5.64 (s, C'3-OH, 0.5 H), 5.30 – 5.18 (m, *J* = 17, 6 Hz, CH=CH_2, 1H), 5.15 – 5.00 (m, *J* = 12, 5 Hz, CH=CH_2, 1H), 4.77 – 4.70 (q, *J* = 8 Hz, CH-OH, 1H), 4.28 – 4.22 (m, C'1-H, 1H), 4.08 – 3.98 (m, C'3-H, 1H), 3.73 – 3.64 (m, CH_2, 2H), 3.50 (m, C'4-H, 1H) 3.20 (dd, *J* = 13, 10 Hz, C'5-H, 1H), 3.12 – 3.04 (dd, *J* = 18, 3 Hz, C'5-H, 1H), 2.92 – 2.86 (dd, *J* = 13, 3 Hz, C'2-H, 1H), 2.87 – 2.81 (dd, *J* = 13, 9 Hz, C'2-H, 1H)

Scheme 5-1: Reaction scheme for EB-Fapy nucleoside adduct formation starting from 2'-deoxyguanosine.

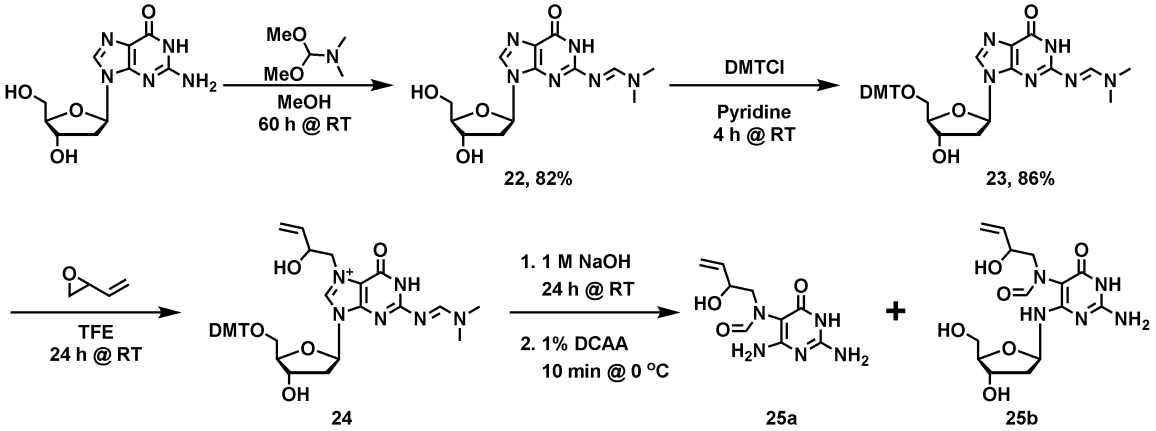
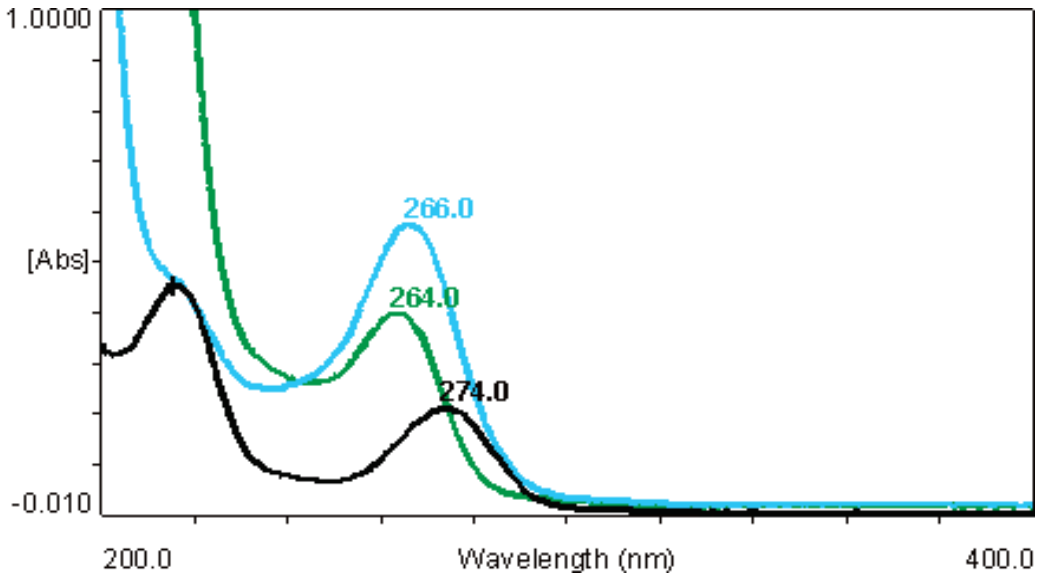


Figure 5-2: UV-visible spectra of EB-Fapy base peaks in solutions of 0.1N HCl (pH: 1.0), water (pH: 7.0) and 0.1N NaOH (pH: 12.0). $\lambda_{max,\ neutral\ pH}$: 266 nm, $\lambda_{max,\ acidic\ pH(1.0)}$: 264 nm, $\lambda_{max,\ basic\ pH(13.0)}$: 274 nm.



UV-Vis characterization of EB-FAPy nucleoside and molar concentration measurement (Figure 5-2):

UV measurements were performed for synthesized EB-Fapy base peaks in solutions of 0.1N HCl (pH: 1.0), water (pH: 7.0) and 0.1N NaOH (pH: 12.0) for the determination of absorption maxima. Molar concentrations of EB-Fapy standard solutions were determined by UV spectrophotometry using the molar extinction coefficient (ε) of 12589.2 M-1 cm^{-1} at 266 nm at neutral pH.[582]

Treatment of Calf-Thymus DNA with 3,4-epoxy-1-butene (EB) to determine dose-dependent increase in EB-FAPy-dG

Calf thymus DNA (1 mg/mL in 10 mM Tris-HCl buffer, pH 7.5) was treated with 10 μM, 50 μM, 100 μM, 250 μM, 500 μM, 1 mM, 5 mM, and 10 mM EB at 37 °C for 24 h. Any unreacted EB was extracted with diethyl ether (3× at 2 mL), and alkylated DNA was precipitated with cold ethanol. The recovered DNA was washed with 70% EtOH and 100% EtOH. DNA was reconstituted in 200 μL 1 N NaOH (pH 12) and kept for 1 h at room temperature to induce EB-Fapy adduct formation. The solution was neutralized with 0.1M HCl, followed by ethanol precipitation of DNA as described above. The DNA was then reconstituted in 200 μL 10 mM tris, pH 7.5 in milli-Q water and DNA was confirmed by nanodrop UV spectrophotometer measurement (Thermo Scientific, Waltham, MA). DNA concentration was determined by dG analysis described in **Section 3.2**.

Enzymatic hydrolysis of DNA and sample preparation for nanoLC-ESI$^+$-MS/MS

DNA (100 μg) was enzymatically digested in the presence of nuclease P_1 (5 U) in 50 μL 10 mM tris-HCl/5 mM $ZnCl_2$ (pH 5.5) for 18 h. The pH of the solution was raised to 7.0 by the addition of 1 μL of 1 M ammonium bicarbonate, followed by incubation with 120 mU phosphodiesterase I, 105 mU phosphodiesterase II, 35 U DNase, and 22 U alkaline phosphatase, in the presence of 15 mM $MgCl_2$ at 37 °C for 18 hours to yield 2'-deoxynucleosides.

The enzymatic digests were spiked with $^{15}N_3$-EB-FAPy-dG internal standard (1 pmol) and purified by solid-phase extraction on Hypercarb Hypersep SPE cartridges (100 mg/3 mL). SPE cartridges were conditioned using 3 mL of methanol, followed by 3 mL of water. DNA samples (200 μL) were loaded under neutral pH, washed with 3 mL of water and 10 % methanol, and then eluted with 70 % methanol in water. SPE fractions were dried under vacuum and reconstituted in 16 μL of 5 mM ammonium formate, pH 4.5 for nanoLC-ESI$^+$-MS/MS analysis as described below.

EB-FAPy-dG and EB-Gua II formation in EB-treated CT DNA

To investigate pH dependence for EB-Fapy-dG and EB-Gua II adduct formation, calf thymus DNA (1 mg/mL in 10 mM Tris-HCl buffer, pH 7.5) was treated with 5 mM EB at 37 °C for 24 h (in duplicate). Following extraction of unreacted EB with diethyl ether (3 x 2 mL), DNA was precipitated with ice cold ethanol. The recovered DNA was washed with 70% ethanol and re-suspended in 10 mM Tris buffer (200 μL) adjusted to pH 7.5, 9.0, 10.0, 11.0, or 12.0. DNA was incubated 1 h, precipitated with cold ethanol, and reconstituted in 10 mM Tris buffer, pH 7.5 (200 μL). Samples were incubated at 70 ºC for

1 hour to release any remaining EB-Gua as a free base. The DNA backbone was precipitated using 100% EtOH and washed as described above, while the supernatant was concentrated under vacuum. Each sample was reconstituted in 200 μL of water, spiked with $^{15}N_5$-EB-Gua II internal standard (200 fmol),[563] and subjected to SPE enrichment using hypercarb SPE cartridges (100 mg, Thermo).

To quantify EB-Fapy-dG adducts, the precipitated, partially depurinated DNA was reconstituted in 10 mM Tris, pH 7.5 buffer (200 μL). DNA aliquots (100 μg) were taken and subjected to enzymatic digestion and solid phase extraction on Hypersep Hypercarb cartridges as described above. EB-Gua II and EB-Fapy-dG adduct numbers were expressed per 10^6 normal nucleotides, and the data were plotted together as a function of pH.

Cytotoxicity of EB in mouse embryonic fibroblasts (MEF) NEIL1$^{+/+}$ and NEIL$^{-/-}$ cells

Wild type and NEIL$^{-/-}$ MEF cells[581] were plated in 100 μL of Dulbecco's Modified Eagle's Medium containing 10% FBS at a density of 1 x 10^4 cells/well. Cells were permitted to adhere in an incubator humidified atmosphere of 5% carbon dioxide, 95% air, at 37 °C for 24 h. Cells were treated with 0 – 20 mM EB in 100 μL of serum-free medium for 3 h at 37 ºC. Following treatment, cell media was replaced with 200 μL fresh FBS-containing media, and the cells were maintained for an additional 48 h at 37 ºC. Cell viability was determined using an Alamar Blue assay[441] using a Synergy HI Microplate reader (BioTek, Winooski, VT).

EB-Fapy-dG and EB-Gua II adduct formation in EB-treated MEF cells

Wild type and NEIL$^{-/-}$ MEF cells were grown in Dulbecco's modified Eagle's media (DMEM) supplemented with 10% fetal bovine serum. Cells were cultured in a humidified atmosphere of 5% carbon dioxide, 95% air, at 37 °C in 15 cm dishes, with at least 70% confluency. Adhered HT1080 cells (1 x 10^7) were treated with 500 μM EB (in triplicate) for 24 hours at 37 °C. Following treatment, EB containing media was discarded, and the cells were washed twice with ice-cold phosphate-buffered saline (PBS) to remove EB. Cells were harvested by trypsin treatment, and DNA was extracted as reported previously using Qiagen Cell Lysis solution.[207] DNA concentrations were determined by dG analysis as described in **Section 3.2**. Extracted DNA (50 μg) was subjected to enzymatic digestion and enrichment for EB-FAPy-dG and EB-Gua II analysis as described above.

EB-Fapy-dG adduct nanoLC-ESI$^+$-MS/MS analysis

Quantitative analyses of EB-FAPy-dG were conducted by nanoLC-ESI$^+$-MS/MS using a TSQ Quantiva mass spectrometer (Thermo Scientific, Waltham, MA) interfaced with a Dionex UltiMate 3000 RSLC nanoHPLC system (Thermo Scientific, Waltham, MA). Samples were loaded onto a pulled-tip fused silica column with a 100 μm inner diameter packed in-house with 15 cm of 5μm Synergy Hydro RP resin (Phenomenex, Torrence, Ca) that served both as a resolving column and as a nanospray ionization emitter. HPLC solvents were comprised of 5 mM ammonium formate, pH 4.5 (solvent A) and 90/10 acetonitrile/solvent A (solvent B). Solvent composition was kept at 2% solvent B over 6 min, followed by an increase to 20% over 10 min, further to 80% over 2 min, held at 80%

B over 2 min, then decreased to 2% over 2 min, and finally re-equilibrated at 2% B for 11 min. The flow rate was 1.0 μL/min for the first 6 min of the run, then reduced to 0.3 μL/min, until the final equilibration for 11 min at 1 μL/min. Under these conditions, EB-FAPy and its internal standard ($^{15}N_3$–EB-FAPy) eluted at ~ 13.8 min. Electrospray ionization was achieved at a spray voltage of 2600 V and a capillary temperature of 350 °C. Collision induced dissociation was performed with Ar as a collision gas (1.0 mTorr) at a collision energy of 10 V. MS/MS parameters were optimized for maximum response during infusion of a standard solution of EB-FAPy.

NanoLC-ESI$^+$-MS/MS analysis of EB-FAPy was performed in the selected reaction monitoring mode by following the neutral loss of dR and both dR and water from protonated molecules of the analyte (*m/z* 356.1 $[M + H]^+ \rightarrow$ 240.1 $[M + H - dR]^+$ and *m/z* 356.1 $[M + H]^+ \rightarrow$ 222.1 $[M + H - dR - H_2O]^+$) (**Figure 5-3A)**. The corresponding mass transitions corresponding to $^{15}N_3$-labeled internal standard were *m/z* 359.1 $[M + H]^+ \rightarrow$ 243.1 $[M + H - dR]^+$ and *m/z* 359.1 $[M + H]^+ \rightarrow$ 225.1 $[M + H - dR - H_2O]^+$) (**Figure 5-3B**). Analyte concentrations were determined using the relative response ratios calculated from HPLC-ESI$^+$-MS/MS peak areas in extracted ion chromatograms corresponding to EB-FAPy and its internal standard.

Figure 5-3A: Observed fragmentation patterns of EB-FAPy-dG by infusion, including the neutral loss of sugar and the loss of sugar and water; 356.1 $[M + H]^+ \rightarrow$ 338.21 $[M - H_2O + H]^+$, 240.1 $[M - dR + H]^+$, 222.1 $[M - dR - H_2O + H]^+$, and 194.1 $[M - dR - H_2O - CO + H]^+$.

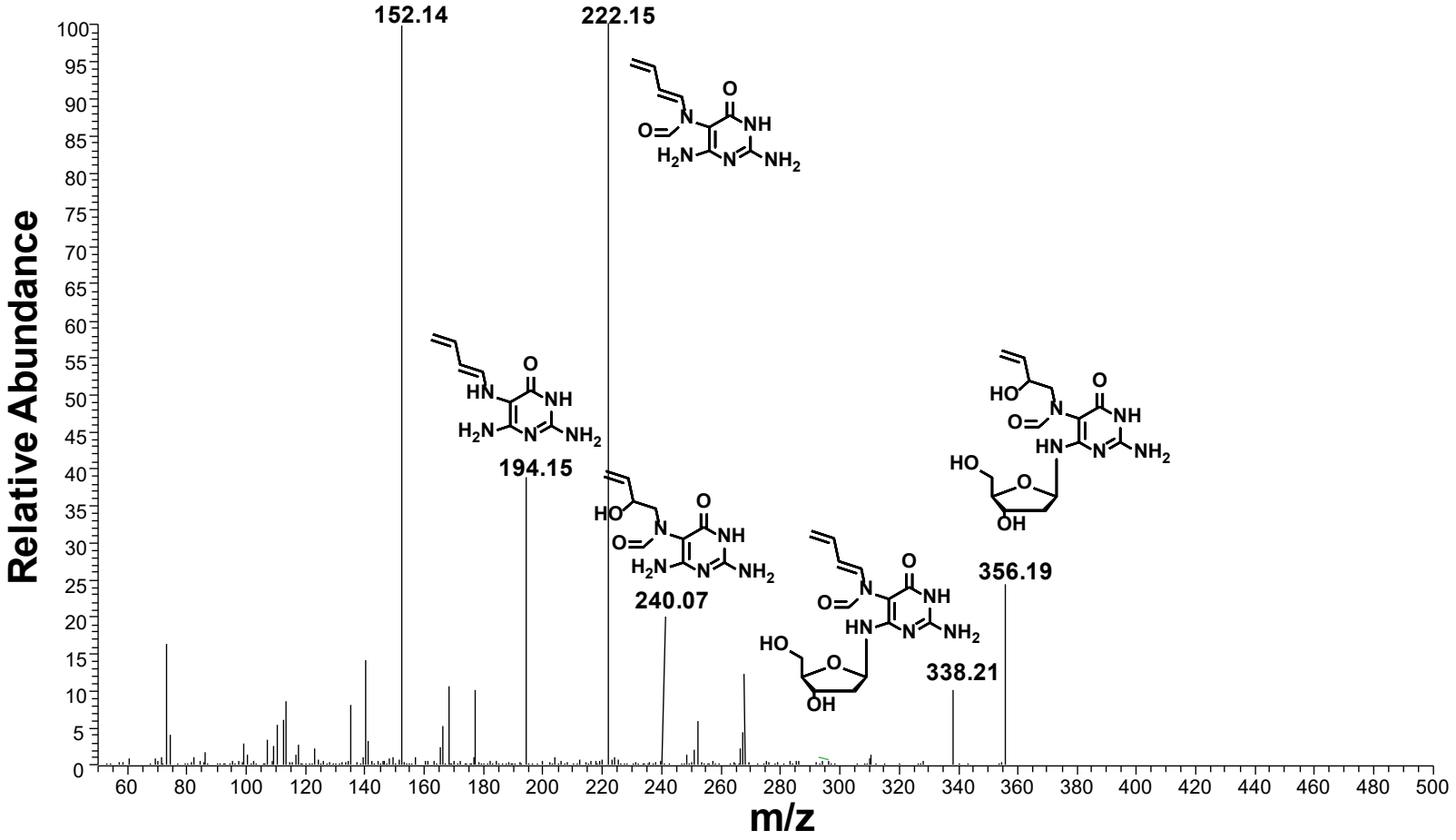
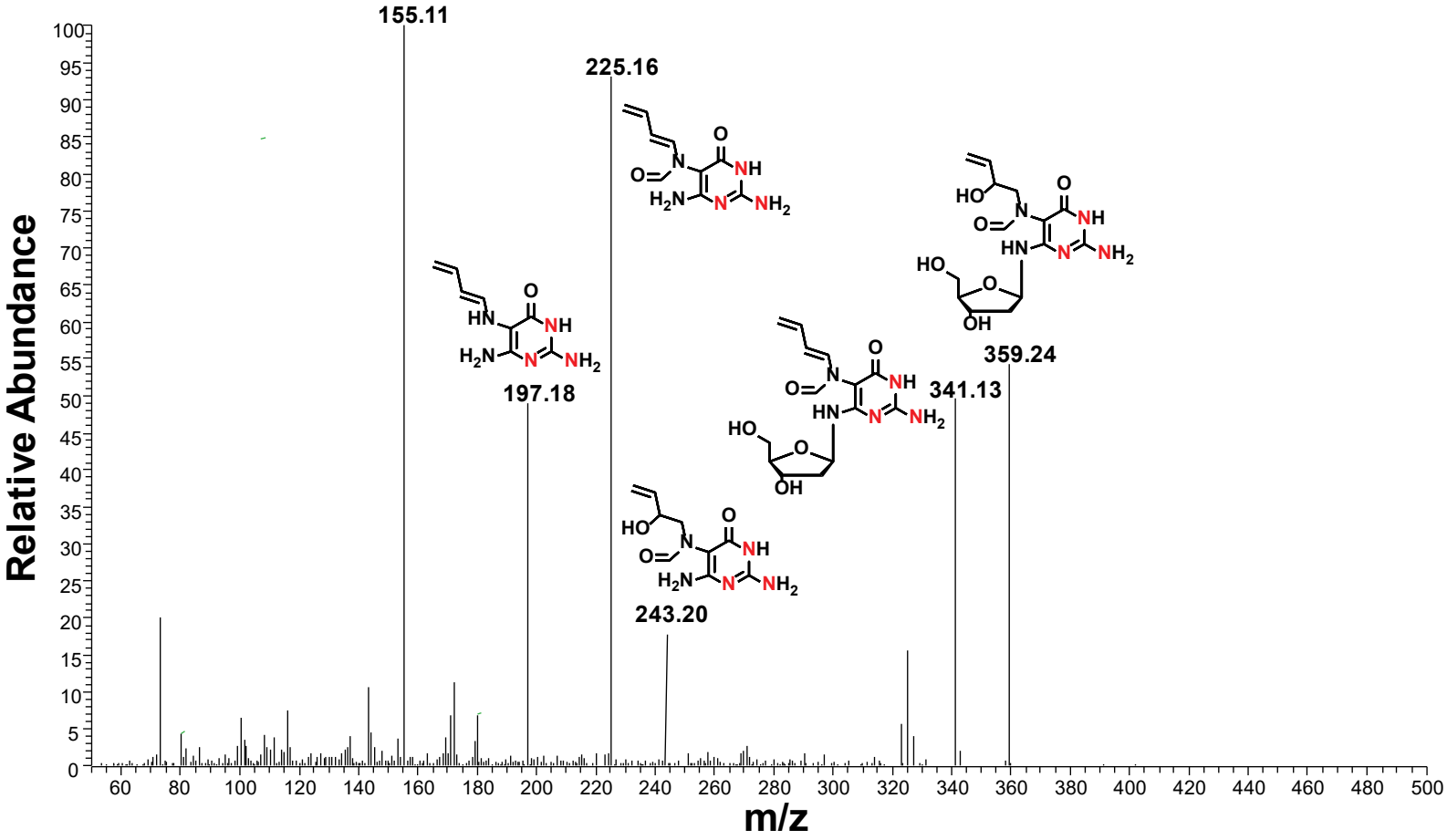


Figure 5-3B: Observed fragmentation patterns of EB-FAPy-dG by infusion, including the neutral loss of sugar and the loss of sugar and water; 359.1 $[M + H]^+ \rightarrow$ 341.21 $[M - H_2O + H]^+$, 243.1 $[M - dR + H]^+$, 225.1 $[M - dR - H_2O + H]^+$, and 197.1 $[M - dR - H_2O - CO + H]^+$.



EB-Gua II adduct nanoLC-ESI$^+$-MS/MS analysis

NanoLC-ESI$^+$-MS/MS was conducted using a TSQ Quantiva mass spectrometer (Thermo Scientific, Waltham, MA) interfaced with a Dionex UltiMate 3000 RSLC nanoHPLC system (Thermo Scientific, Waltham, MA). Samples were loaded onto a pulled-tip fused silica column with a 100 µm inner diameter packed in-house with 15 cm of 5µm Synergy Hydro RP resin (Phenomenex, Torrence, Ca) that served both as a resolving column and as a nanospray ionization emitter. HPLC solvents included (A) 0.01% acetic acid in LC-MS grade water and (B) 0.02% acetic acid in LC-MS grade acetonitrile. Solvent composition was kept at 2% solvent B over 8 min, followed by an increase to 30% over 9 min, further to 80% over 1 min, held at 80% B over 10 min, then decreased to 2% over 1 min, and finally re-equilibrated at 2% B for 8 min. The flow rate was 0.7 µL/min for the first 7.5 min of the run, then reduced to 0.3 µL/min, until the final equilibration for 11 min at 0.7 µL/min. Sample injection volume was 4 µL. Under these conditions, both analyte its internal standard eluted around 16.3 min.

Electrospray ionization was achieved at a spray voltage 3100 V and capillary temperature of 350 °C. Collision induced dissociation was performed with Ar as a collision gas (1.5 mTorr). MS tuning was performed upon infusion of EB-Gua II internal standard to obtain maximum sensitivity. HPLC-ESI$^+$-MS/MS analysis was performed in the selected reaction monitoring (SRM) mode using $^{15}N_5$ labeled internal standard ($^{15}N_5$-EB-Gua II). SRM transitions included neutral loss of 1-hydroxy-3-buten-2-yl from protonated EB-Gua II analyte *m/z* 222.1 $(M + H)^+ \rightarrow$ *m/z* 152.1 $(M - C_4H_8O + H)^+$ with a collision energy (CE) of 15 V and a neutral loss of 1-hydroxy-3-buten-2-yl and ammonia from protonated EB-Gua II (*m/z* 222.1 $(M + H)^+ \rightarrow$ *m/z* 135.0 $(M - C_4H_8O - NH_3 + H)^+$) with a CE of 29 V.

Similarly, SRM transitions for $^{15}N_5$-EB-Gua II included *m/z* 227.1 ($^{15}N_5$-M + H)$^+$ →*m/z* 157.0 ($^{15}N_5$-M – C_4H_8O + H)$^+$, CE = 17 V and *m/z* 227.1 ($^{15}N_5$-M + H)$^+$ → *m/z* 139.0 ($^{15}N_5$-M – C_4H_8O – $^{15}NH_3$ + H)$^+$, CE = 32 V. Quantitative analyses was performed on the ratio of the area under the peak using SRM transitions *m/z* 222.1 (M + H)$^+$ → *m/z* 152.1 (M – C_4H_8O + H)$^+$ for the analyte and *m/z* 227.1 ($^{15}N_5$-M + H)$^+$ → *m/z* 157.0 ($^{15}N_5$-M – C_4H_8O + H)$^+$ for the internal standard.

NanoLC-ESI$^+$-MS/MS method validation

Calf thymus DNA (CT DNA, 100 μg, in triplicate) was spiked with increasing amounts of EB-FAPy-dG analyte (40-2000 fmol) and $^{15}N_3$-EB-FAPy-dG internal standard (1 pmol). Each sample was processed as described above. The linearity of our nanoLC-ESI$^+$-MS/MS methodology was determined by plotting the observed analyte/internal standard peak area ratio versus the expected peak area ratio. The limit of detection was defined as the analyte signal at least 3 times the height as compared to the background. Accuracy and precision of the method was calculated for three samples (1000 fmol, high point; 400 fmol, middle point; 100 fmol, low point). Accuracy was calculated by subtracting the average observed peak area ratio (analyte/internal standard) from the expected peak area ratio, followed by dividing the difference by the expected peak area ratio multiplied by 100. Accuracy was expressed as the % difference. Precision was calculated by dividing the standard deviation between samples by the mean observed peak area multiplied by 100, expressed as 100 - % coefficient of variation (100 - %CV). Validation results are presented below in the results.

5.3. Results

5.3.1. Synthesis and structural characterization of EB-FAPy-dG adduct

In our initial efforts to synthesize authentic standards of EB-Fapy-dG adducts, 2'-deoxyguanosine (dG) was reacted with excess *S,R*-EB at pH 7.2, followed by base treatment with 1N NaOH to induce imidazole ring opening reaction. Although ESI^+-MS/MS analyses have revealed the presence of small amounts of EB-Fapy-dG (M = 356.1), they could not be purified due to the low adduct yields and the complexity of the reaction mixtures (results not shown).

To improve EB-Fapy-dG yields, doubly protected dG (**23** in **Scheme 5-1**) was employed as a starting material. This approach was previously employed by Rizzo *et al* in the synthesis of Me-FAPy-dG phosphoramidite.[231] Following reaction with EB to generate N7-alkylated dG **24**, the reaction mixtures were incubated in the presence of 1N NaOH to induce imidazole ring opening. Following neutralization and deprotection steps, reaction mixtures were separated by semi-preparative reverse phase HPLC. All major HPLC peaks were collected and subjected to HPLC-ESI^+-MS/MS analysis. We found that the peaks eluting at 9.5, 10.7, and 11.8 min contained EB-FAPy as a free base (M = 240.1), while the HPLC fractions collected at 13.7, 14.1, and 14.8 min contained EB-FAPy nucleoside (M = 356.1).

HPLC purified EB-FAPy base and EB-FAPy-dG were first structurally characterized by MS infusion. The EB-FAPy base was characterized by the loss of water and the loss of water and carbon monoxide from the protonated parent molecules (*m/z* 240.1 $[M + H]^+ \rightarrow$ 222.1 $[M - H_2O + H]^+$ and 194.1 $[M - H_2O - CO + H]^+$. EB-FAPy-dG was characterized by a loss of sugar, a loss of sugar and water, and a loss of sugar, water,

and carbon monoxide from the protonated parent molecules (*m/z* 356.1 $[M + H]^+ \rightarrow$ 240.1 $[M + H - dR]^+$, 222.1 $[M + H - dR - H_2O]^+$, and 194.1 $[M - dR - H_2O - CO + H]^+$ (**Figure 5-3A**).

EB-FAPy base and EB-FAPy nucleoside (Compound **25a** and **25b** in **Scheme 5-1**) were further structurally characterized by ^{1}H-NMR and ^{1}H-^{13}C HSQC using a 600 MHz NMR. ^{1}H-NMR confirmed the presence of the 2-hydroxybut-3-en-1-yl group as revealed by the observation of geminal alkene protons at 5.22 and 5.08 ppm, internal alkene at 5.82 ppm, a quartet at 4.26 ppm (2-position of substituent), and a multiplet at 3.51 ppm (**Figure 5-4A**). Further analysis of the EB-FAPy base by ^{1}H-^{13}C HSQC spectroscopy revealed ^{1}H signals at 5.80, 5.22, and 5.08 ppm, which correlated to ^{13}C signals at 117.6 and 134.5, confirming the presence of a terminal alkene (**Figure 5-4B**). Furthermore, the ^{1}H-quartet at 4.74 ppm and a multiplet at 3.51 ppm correlated with carbon atoms observed at 63.2 and 60.9 ppm, respectively, as expected for the 2-hydroxybut-3-en-1-yl substituent (**Figure 5-4B**).

The presence of formamido group was confirmed by the presence of ^{1}H signals at 7.78 and 8.10 ppm, which correlated to ^{13}C signals at 164.2 and 166.2 ppm in the HSQC analysis (**Figure 5-4B**). Our observation of two major formamide peaks is consistent with the EB-FAPy molecule existing as a mixture of atropisomers caused by sterically hindered rotation about the C5-N^5-bond (**Figure 5-4A**).[224, 230, 231] Additional splitting of the formamide peaks can be attributed to the presence of another pair of rotamers caused by the slow rotation around the N^5-CO bond respectively (**Figure 5-4B**). *N*1H, exocyclic amines at C2 and C4, and the alcohol OH were observed at 10.34, 6.47, 6.34, and 6.13 ppm respectively (**Figure 5-4A)**.

^{1}H-NMR and ^{1}H-^{13}C HSQC analysis of EB-FAPy nucleoside confirmed its structure as *N*^{6}-(2-Deoxy-D-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-*N*-(2-hydroxybut-3-en-1-yl)formamidopyrimidine (**Figure 5-5A**). Additional peaks corresponding to 2'-deoxyribose were observed at 4.25, 4.03, 3.50, 3.20, 3.08, 2.89, and 2.84 ppm. Compared to the NMR spectrum of the free EB-FAPy base, EB-FAPy nucleoside exhibited a small downfield shift of the protons corresponding to 2-hydroxybut-3-en-1-yl substituent observed at 4.74 ppm and 3.68 ppm (**Figure 5-5A**). Furthermore, ^{1}H signals corresponding to the formamide group were observed at 7.90, 7.83, and 7.77 ppm. However, HSQC spectra revealed no significant changes in ^{13}C chemical shifts corresponding to the 2-hydroxybut-3-en-1-yl substituent (**Figure 5-5B**). As was the case for NMR spectra of EB-FAPy base (**Figure 5-4A**), multiple signals corresponding to the formamide protons were observed due to the presence of multiple atropisomers and rotamers of the adduct (**Figure 5-5B**).*[224, 230, 231]*

Collectively, our mass spectrometry and NMR results are consistent with the structure of *N*^{6}-(2-Deoxy-D-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-*N*-(2-hydroxybut-3-en-1-yl)formamidopyrimidine (EB-Fapy-dG, 4 in Scheme 2). Isotopically labeled analog, $^{15}N_3$- EB-FAPy-dG, was synthesized analogously using $^{15}N_3$-dG as a starting material.

Figure 5-4: H-NMR (**A**) and HSQC (**B**) analysis of EB-FAPy base (**25A**)

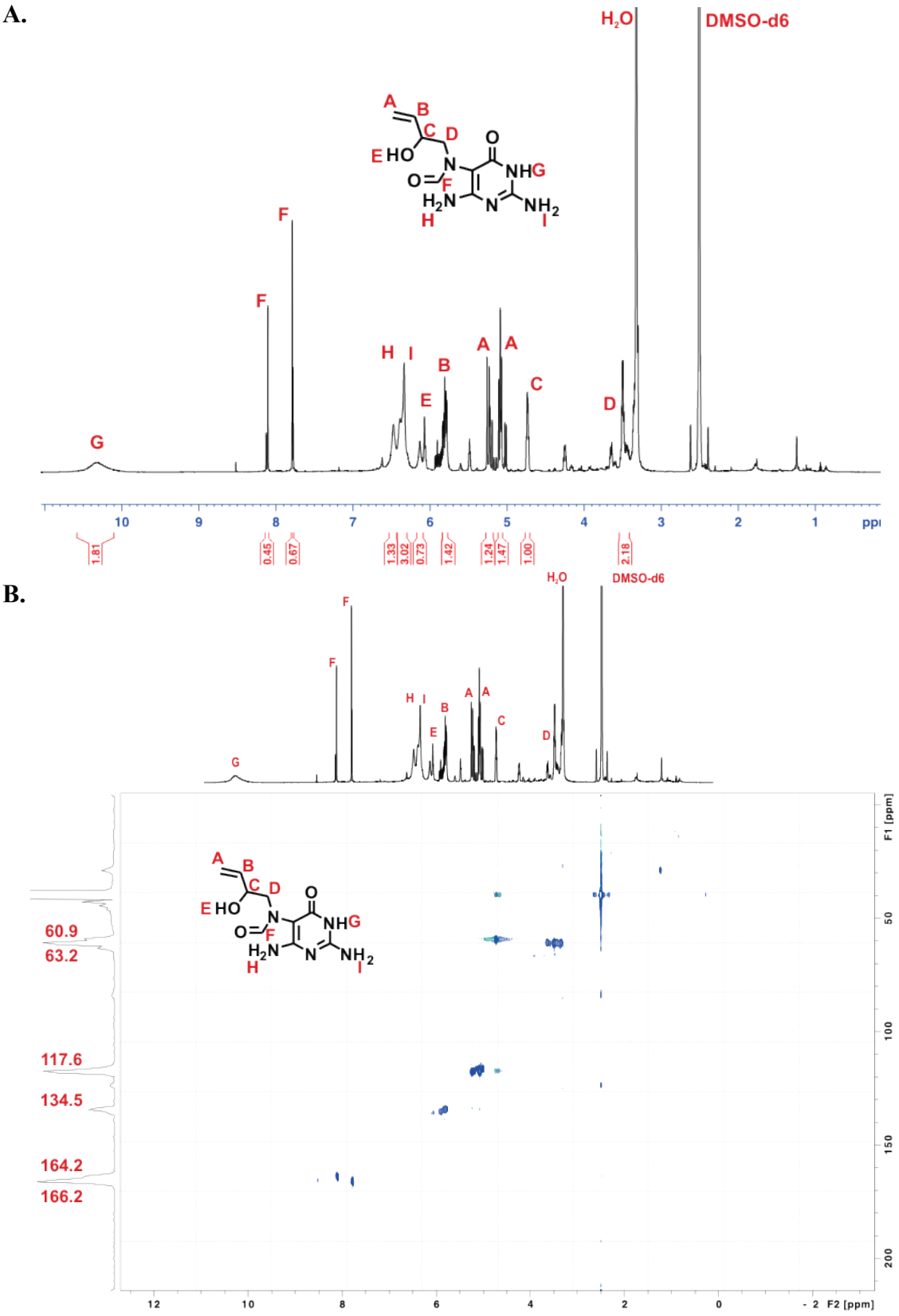
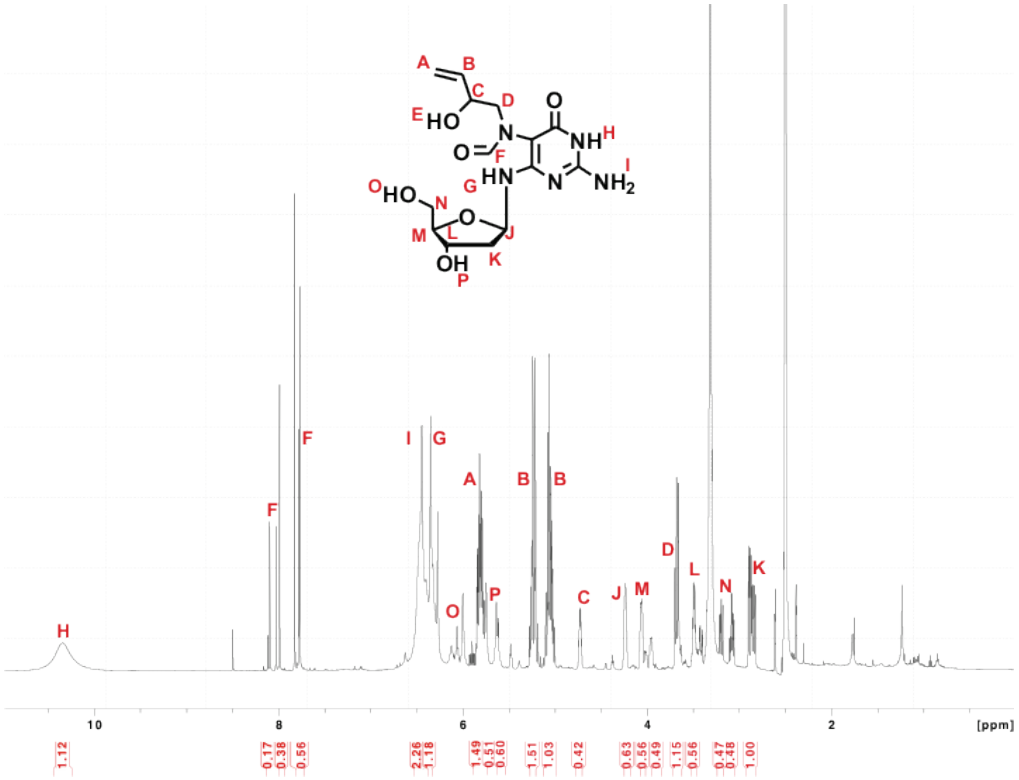
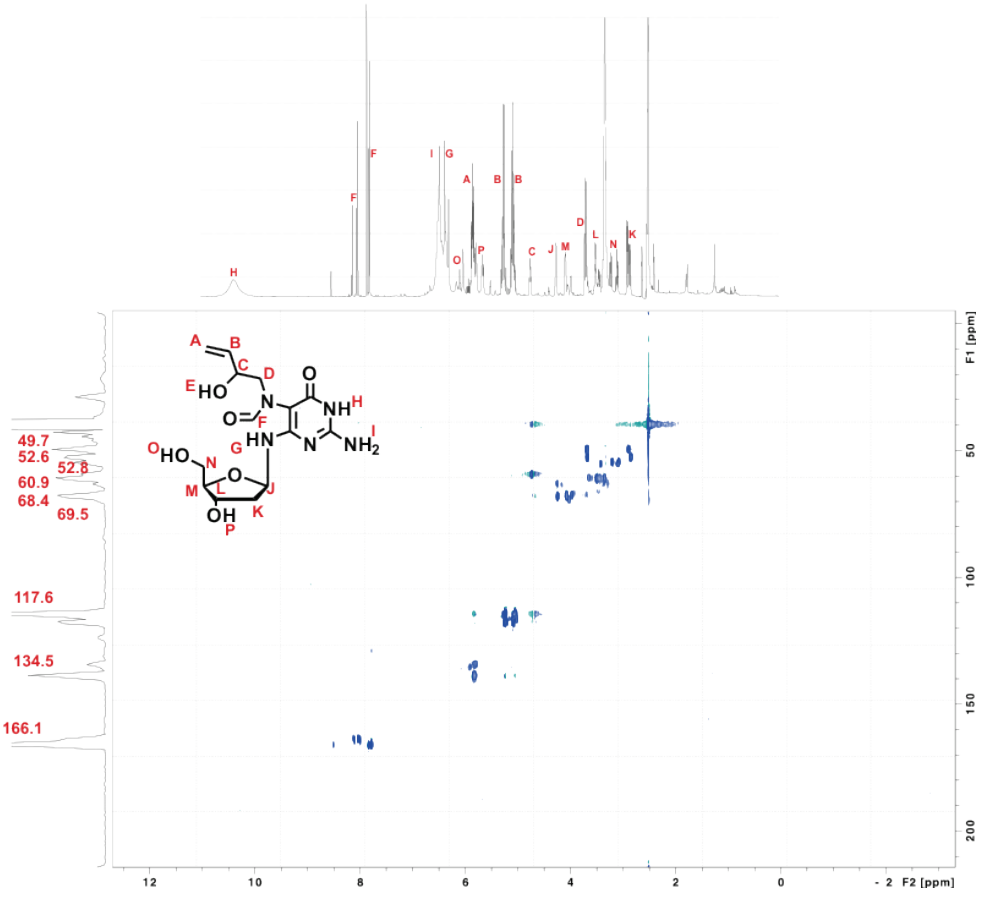


Figure 5-5: H-NMR (**A**) and HSQC (**B**) analysis of EB-FAPy nucleoside (**25B**)

A.



B.



5.3.2. nanoLC-ESI$^+$-MS/MS method development and validation

Authentic standards of EB-FAPy-dG and its $^{15}N_3$- labeled isotopomer ($^{15}N_3$- EB-FAPy-dG) were used to develop an isotope dilution nanoLC-ESI$^+$-MS/MS method for accurate quantitation of the adduct in EB-treated DNA and in human cells exposed to EB. In our approach, DNA (50 μg) was fortified with known amounts of $^{15}N_3$- EB-FAPy-dG internal standard, followed by enzymatic hydrolysis to release EB-FAPy-dG nucleoside. EB-FAPy-dG and its internal standard are enriched by solid phase extraction on Carbograph SPE cartridges. BD induced EB-Fapy adducts are very polar molecules and do not retain well on SPE columns. We found that among SPE cartridges tested (Isolute ENV+ (50 mg/1 mL, Biotage, Charlotte, NC), Sep-pack C18 (50 mg/1 mL, Waters Corp., Millford, MA), Strata X polymeric C18 (30 mg/ 1mL, Phenomenex, Torrance, CA)) graphitised carbon based packing (Carbograph, Grace or Hypersep Hypercarb, ThermoFisher Scientific) was the only one that allowed for good retention of EB-FAPy-dG, leading to good SPE recovery (90.3%) and the removal of impurities that interfere with nanoLC-ESI$^+$-MS/MS.

SPE fractions containing EB-FAPy-dG were concentrated under vacuum and re-constituted for isotope dilution nanoLC-ESI$^+$-MS/MS analysis. We chose nanoLC because it allows for the best sensitivity in mass-spectrometry based approaches for DNA adduct detection.[202] Because of its highly polar nature (calculated lopP = – 1.498), EB-FAPy-dG does not retain well on standard C18 packing materials. Furthermore, it exhibits a complex HPLC peak shape due to the presence of multiple rotamers/ atropisomers in its tructure. A number of HPLC solid phases (Synergi Max-RP (Phenomenex, Torrance, CA), Hypersil Gold aq (Thermo Scientific, Waltham, MA), Zorbax SB-C18 (Agilent technologies, Santa

Clara, CA) and solvent systems (0.05% acetic acid, 0.1 % formic acid, 15 mM ammonium acetate (pH 6.8) with methanol or acetonitrile) were tested. The best HPLC retention and peak shape were achieved with Synergi Hydro RP packing material (Phenomenex, Waltham, MA) (**Figure 5-6**). When using a gradient of acetonitrile in 5 mM ammonium formate, pH 4.5, EB-FAPy-dG eluted as a complex peak at 12.5-14 min (**Figure 5-6**). Our attempts to focus all EB-FAPy-dG rotamers into a single peak were unsuccessful as they compromised analyte retention and reduced the sensitivity.

As noted above, MS/MS spectrum of EB-Fapy-dG (*m/z* 356.1, **Figure 5-2A**) revealed two major fragments corresponding to the neutral loss of deoxyribose sugar (*m/z* 240.1) and both deoxyribose sugar and a molecule of water (*m/z* 222.1). $^{15}N_3$-EB-Fapy-dG exhibited analogous signals at *m/z* 243.1 and 225, respectively (**Figure 5-2B**). Therefore, our quantitative nanoLC-ESI$^+$-MS/MS methodology is based on selected reaction monitoring of mass transitions *m/z* 356.1 → 240.1 (analyte) and *m/z* 359.1 → 243.1 (internal standard). Secondary (qualifier) SRM transitions (*m/z* 356.1 → 222.1 (analyte) and *m/z* 359.1 → 225.1 (internal standard) were used to confirm signal identities (**Figure 5-2B**). Mass spectrometer parameters were optimized for maximum sensitivity upon infusion of synthesized standards.

The newly developed nanoLC-ESI$^+$-MS/MS method was validated by analyzing CT DNA samples spiked with known amounts of the analyte (40 – 2000 fmol) and 1000 fmol internal standard. The nanoLC-ESI$^+$-MS/MS limit of detection (LOD) was 10 fmol EB-FAPy-dG on column, while the limit of quantification was 25 fmol EB-FAPy-dG on column (**Figure 5-7**). The accuracy (% difference) of the method was calculated as 97.7%, 91.7%, and 97.8% for the high, middle, and low validation points. The accuracy (% CV)

of the method was calculated as 0.3, 3.9, and 3.1 % CV for the high, middle, and low validation points.

Figure 5-6: NanoLC-ESI$^+$-MS/MS analysis chromatogram of EB-FAPy-dG standard (200 fmol on column) and $^{15}N_3$-internal standard (200 fmol on column) in water showing two different MS/MS transitions of the protonated parent ion.

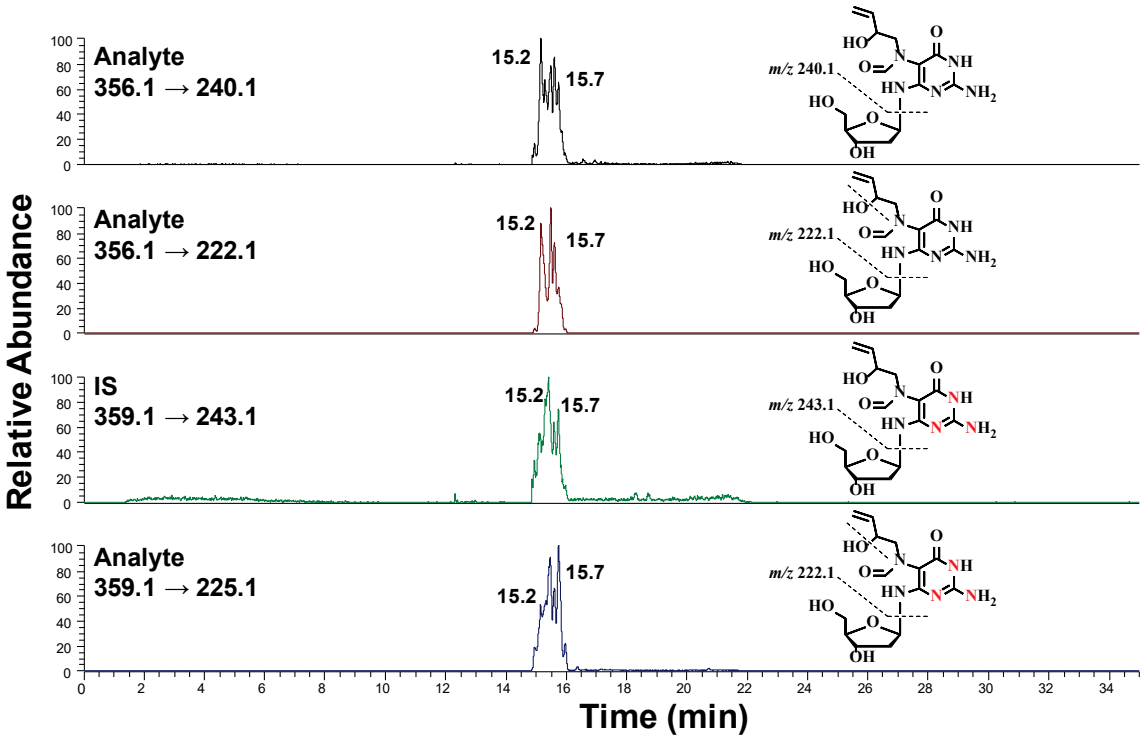
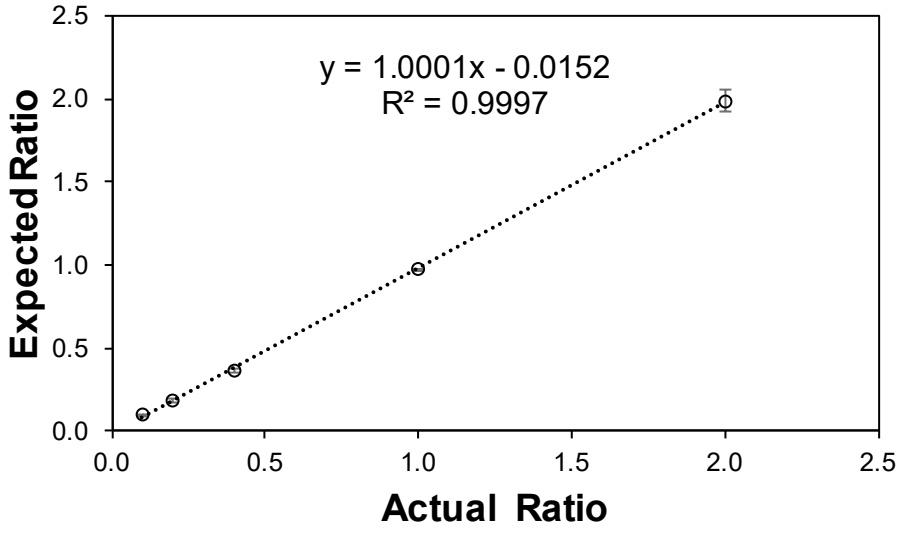


Figure 5-7: Validation of the developed EB-FAPy-dG nanoLC-ESI$^+$-MS/MS assay. Calf thymus DNA (50 µg, in triplicate) was spiked with EB-FAPy-dG analyte (2000, 1000, 400, 200, 100, and 40 fmol) and 1000 fmol $^{15}N_3$-EB-FAPy-dG IS. Each sample was processed for EB-FAPy SPE enrichment as described in the Materials and Methods. The linearity of each validation replicate was determined by plotting the observed peak area ratio (analyte/internal standard) to the expected peak area ratio. The limit of detection was defined as the analyte signal at least 3 times the response as compared to the background.



5.3.3. EB-FAPy-dG adduct detection and quantitation in CT DNA

To determine whether EB-Fapy-dG adducts can be formed in DNA, calf thymus DNA was treated with increasing concentrations of EB. Aliquots of treated DNA were taken and either analyzed directly or pre-treated with NaOH to induce FAPy adduct formation. The new isotope dilution nanoLC-ESI$^+$-MS/MS methodology described above was employed to allow for accurate quantification. Samples pre-treated with NaOH exhibited concentration dependent formation of EB-Fapy-dG (**Figure 5-8**), with adduct numbers ranging from 1.8 to 234.5 EB-FAPy-dG adducts per 10^5 nucleotides, depending on EB concentrations. EB-FAPy-dG adduct numbers increased linearly between 10 μM to 5 mM EB, followed by plateauing at 10 mM (**Figure 5-8**). In contrast, samples not pre-treated with base contained much lower adduct numbers (e.g. 2.09 ± 0.07 EB-FAPy-dG per 10^6 nucleotides for sample treated with 5 mM EB) (**Figure 5-9**).

Figure 5-8: Concentration dependent formation of EB-Fapy-dG adducts in CT DNA treated with increasing concentrations of EB, followed by treatment with base (1N NaOH, 1 h).

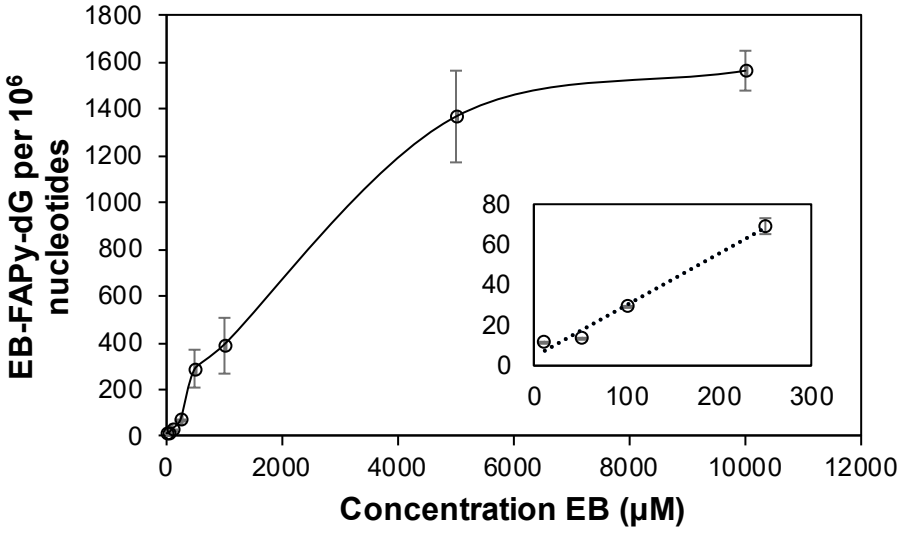
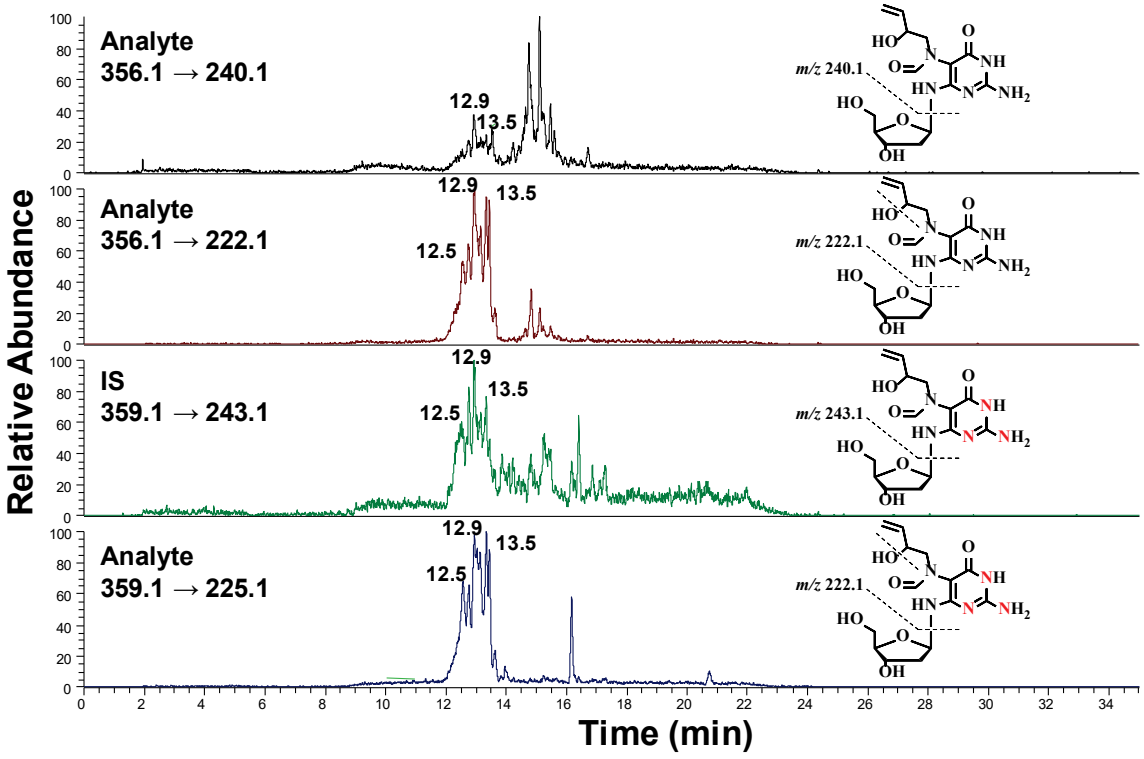
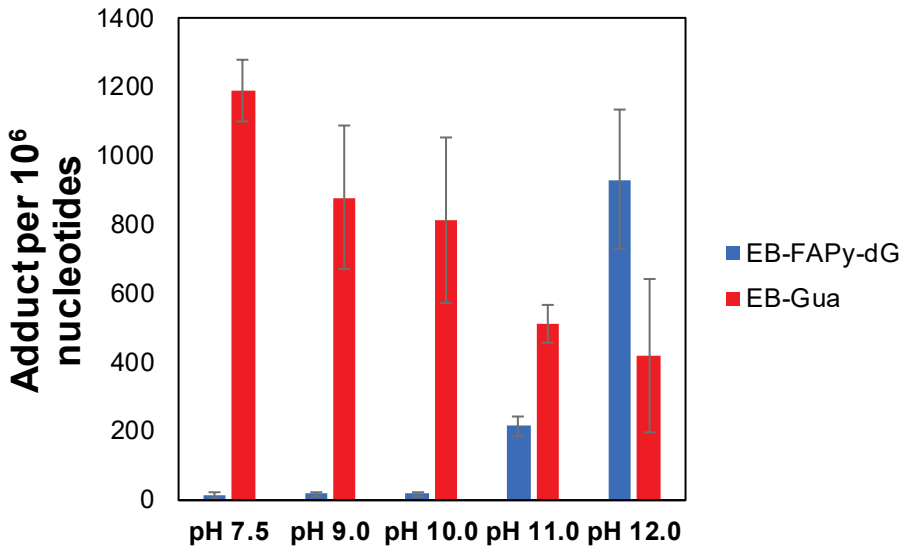


Figure 5-9: Representative trace of EB-FAPy-dG detected in CTDNA treated with 5 mM EB but not treated with 1 M NaOH.



To establish pH requirements for EB-FAPy formation following treatment with EB, and to determine what fraction of N7-EB-dG adducts converts to EB-FAPy, we have also quantified EB-Gua II as a free base using the previously reported methodology.[200, 206, 563] After treating CT DNA with 5 mM EB, the DNA was incubated at increasing pH (pH 7.5, 8.0, 9.0, 10.0, 11.0, 12.0) to induce EB-FAPy formation, followed by heating at 37 ºC for three days to release EB-Gua II as a free base. EB-Fapy adducts were detected in every CT-DNA sample, with levels measured at 20.9 ± 9.7, 26.1 ± 6.3, 27.4 ± 10.5, 322.7 ± 43.7, and 1397.1 ± 306.4 adducts per 10^6 nucleotides at pH 7.5, 9.0, 10.0, 11.0, and 12.0 (**Figure 5-10**). EB-Gua II adducts revealed an inverse relationship to EB-FAPy-dG formation. EB-Gua II levels in the same samples were measured at 1190.2 ± 89.3, 879.4 ± 207.6, 813.8 ± 240.4, 512.3 ± 53.3, and 420.6 ± 223.5 adducts per 10^6 nucleotides at pH 7.5, 9.0, 10.0, 11.0, and 12.0 respectively (**Figure 5-10**).

Figure 5-10: pH dependent EB-FAPy-dG and EB-Gua II adduct formation in CT DNA treated with 5mM EB, incubated at 37 ºC for 3 days to allow EB-FAPy-dG formation, and then heated at 70 ºC for 1 h to release remaining alkylated dG.



5.3.4. NEIL1$^{+/+}$ and NEIL1$^{-/-}$ MEF EB cytotoxicity and EB-FAPy-dG formation in MEF cells

As mentioned above, NEIL1 enzyme has been shown to repair of alkyl-FAPy-dG adducts such as those induced by Aflatoxin B_1 (AFB_1-FAPy-dG) and methylating agents (Me-FAPy-dG).[573, 583] To elucidate the role NEIL1 plays in the repair of EB-FAPy-dG, we first performed cell cytotoxicity experiment to determine the toxicity of EB in wild type and NEIL1$^{-/-}$ MEF cells. Both cell types (in quadruplet) were treated with increasing amounts of EB from 0 to 20,000 μM for 3 hours, followed by incubating an additional 48 hours and measuring cell viability using the Alamar Blue assay.[441] NEIL1$^{+/+}$ cells were found to be 1.5-fold less sensitive to EB (IC_{50} = 1456 ± 126 μM) than NEIL1$^{-/-}$ cells (IC_{50} = 1006 ± 108 μM) (**Figure 5-11**).

We next investigated EB-Fapy-dG and EB-Gua II adduct formation in both NEIL1$^{+/+}$ and NEIL1$^{-/-}$ MEF cells treated with 500 μM of EB for 24 hours. After enzymatic digestion and adduct enrichment, both EB-FAPy-dG and EB-Gua II were quantified by LC-MS/MS analysis (**Figure 5-12**). NEIL1$^{-/-}$ cells had nearly 2-fold greater numbers of EB-FAPy-dG adducts (6.68 ± 0.51 EB-FAPy-dG per 10^6 nucleotides) compared to NEIL1$^{+/+}$ cells (3.40 ± 0.032 EB-FAPy-dG per 10^6 nucleotides, **Figure 5-12**). Similarly, NEIL1$^{-/-}$ cells had nearly 3-fold greater EB-Gua II adducts (14.01 ± 1.52 EB-Gua II per 10^6 nucleotides) compared to NEIL1$^{+/+}$ cells (5.35 ± 0.22 EB-Gua II per 10^6 nucleotides, **Figure 5-12**).

Figure 5-11: Cell cytotoxicity to EB treatment in NEIL1$^{+/+}$ and NEIL1$^{-/-}$ MEF cells.

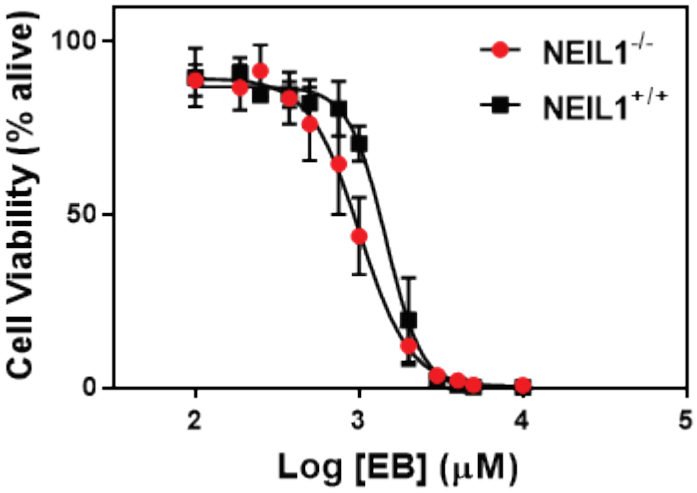
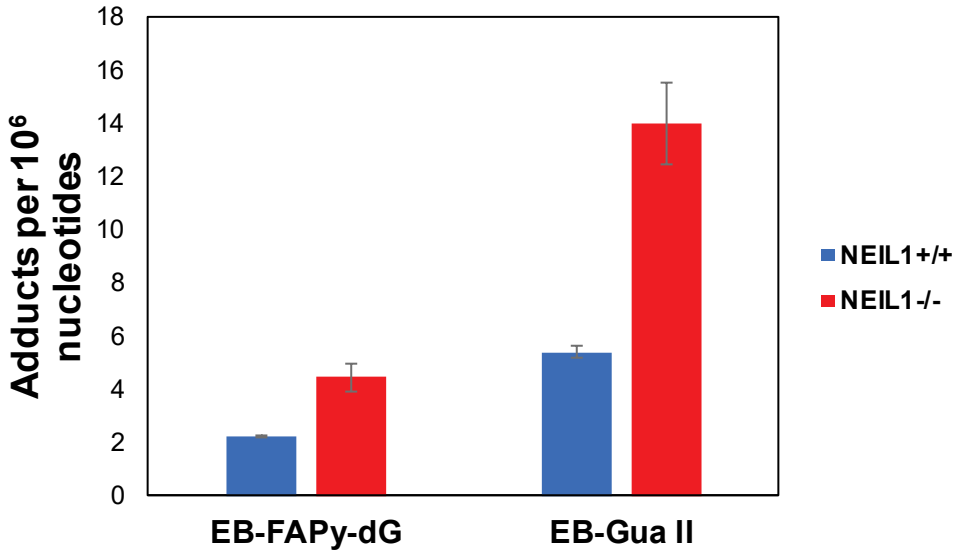


Figure 5-12: Quantitation of EB-FAPy-dG and EB-Gua adducts in NEIL1$^{+/+}$ and NEIL1$^{-/-}$ MEF cells treated with 500 μM EB. NEIL1$^{-/-}$ cells had nearly 2-fold greater numbers of EB-FAPy-dG adducts (4.45 ± 0.51 compared to 2.27 ± 0.032 EB-FAPy-dG per 10^6 nucleotides, p-value < 0.01) and nearly 3-fold greater EB-Gua II adducts (14.01 ± 1.52 compared to 5.35 ± 0.22 EB-Gua II per 10^6 nucleotides, p-value < 0.01).



5.4. Discussion

1,3-butadiene (BD) is a known carcinogen present in cigarette smoke,[556] automobile exhaust,[177] urban air, and still utilized in industrial processes.[184] Previous studies have confirmed the reactive epoxide metabolites of BD, such as EB, DEB, and EDB, are all capable of alkylating nucleophilic positions of DNA.[202, 558] The formation of covalent BD-DNA adducts are thought to be responsible for BD-induced carcinogenicity, as these adducts can lead to mutations during DNA replication.[196, 197, 559-562] However, cationic N7-purinic adducts are not believed to be mutagenic because the N7-position of purines does not participate in Watson-Crick base pairing.[252, 568] Furthermore, N7-adducts are susceptible to depurination, leaving apurinic sites which are an ubiquitous endogenous lesion that can be efficiently recognized and repaired.[584-586] However, N7-adducts can undergo base-catalyzed imidazole ring opening to yield the hydrolytically stable ans strongly genotoxic N^5-alkyl-Fapy-dG adducts. To our knowledge, no BD-derived FAPy adducts have been reported previously. Furthermore, it is unknown whether these adducts contribute to the observed mutagenic and genotoxic effects of BD.

Authentic N^6-(2-Deoxy-D-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-*N*-(2-hydroxybut-3-en-1-yl)formamidopyrimidine (EB-Fapy-dG) standard was prepared by a synthetic strategy analogous to published methods for other alkyl-FAPy adducts,[231, 232] EB-FAPy free base and EB-FAPy-dG were characterized by UV-vis, ESI^+-MS/MS, ^{1}H-NMR, and ^{1}H-^{13}C HSQC. NMR analysis of purified EB-FAPy revealed this adduct spontaneously undergoes isomerization in solution (**Figure 5-1**). Previous studies with AFB_1-FAPy-dG and Me-FAPy-dG have shown the C5-N^5 bond is capable of restricted rotation to form atropisomers, while the formyl N-CO bond is capable of rotation

as well to yield stable rotamers.[224, 230, 231] Additional isomers include β- and α-anomers and regioisomers formed after ring-opening of the deoxyribose sugar and reclosure of the deoxyribose sugar involving the 5'-hydoxyl group to yield the pyranose nucleoside (**Figure 5-1**).[224, 587] Acid cleavage of the 5'-hydroxyl trityl-protecting group has been shown to rapidly catalyze anomerization and pyranose/furanose conversion in AFB_1-FAPy-dG and Me-FAPy-dG containing DNA.[230, 231] However, we did not observe additional peaks corresponding to the pyranose sugar, suggesting that our method's trityl-deprotection step (single exposure, 10 min at 0 ºC) was not sufficient to stimulate acid-catalyzed deoxyribose ring opening required for nucleoside anomerization.

Our investigation to find the optimal pH for imidazole ring opening of EB-N7-dG revealed that EB-FAPy-dG form at physiological pH, but the optimal pH for adduct formation is pH 12.0 (**Figure 5-10**). EB-FAPy-dG adduct levels in CT DNA treated with 5 mM EB at pH 7.5 – 10.0 were below 30 adducts per 10^6 nucleotides, but significantly increased to 322.7 ± 43.7, and 1397.1 ± 306.4 adducts per 10^6 nucleotides at pH 11.0 and 12.0 respectively. After EB alkylates the N7-position of dG, a positive charge is formed leading to the instability of the N-glycosidic bond and hydrolysis of the modified base from the DNA backbone. EB-Fapy-dG adduct formation competes with this spontaneous hydrolysis reaction, with imidazole ring opening greatly accelerated in presence of free base (OH^-).

We found that incubation at 37 ºC for 72 h does not successfully release EB-N7-dG adducts from DNA backbone, allowing EB-FAPy-dG adducts to form. When we quantified EB-Gua II released from the DNA backbone after this heating step, we detected less EB-Gua II compared to EB-FAPy-dG at every pH analyzed (**Figure 5-13A**).

Furthermore, we were unable to detect EB-FAPy base adduct formation when EB-treated DNA was not further incubated after treatment at pH 7.5 and 9.0 for only 1 hour (**Figure 5-13B**). Previous *in-vitro* studies have shown that N^5-alkyl-FAPy-dG adducts are hydrolytically stable at physiological pH, but significantly less stable under slightly acidic conditions. Christov *et al* observed complete depurination of methyl-FAPy-dG from a 13-mer oligodeoxynucleotide after 48 h at pH 6.5, while the same lesion was stable at physiological pH.[231] Therefore, it is conceivable that any formed EB-FAPy-dG would remain on DNA during our experiment, reaching detectable levels after three days of incubation.

Failure to repair alkyl-FAPy-dG adducts is known to lead to significant biological consequences such as mutagenicity and genotoxicity. Methyl-FAPy-dG adducts have been identified as an inhibitor to DNA replication.[227, 236, 237, 588] Ide *et al* investigated the bypass of a site-specific methyl-FAPy-dG adduct in an oligonucleotide by the Klenow fragment of *E. Coli* DNA polymerase 1 (KF), and observed the FAPy lesion inhibited the extension step of lesion bypass.[237] Similarily, Rizzo *et al* investigated the bypass of methyl-FAPy-dG by eukaryotic DNA polymeraeses α, β, and hPol δ/ PCNA, and observed the lesion blocked all three high fidelity polymerases at either the insertion or extension site.[234] Rizzo *et al* also investigated the human translesion synthesis polymerases hPol η, κ, and hRev1/Pol ζ ability to bypass the lesion.[234] All three TLS polymerases primarily inserted the correct base (2'-deoxycytidine), but hPol η and κ incorrectly inserted thymidine, 2'-deoxyguanosine, and 2'-deoxyadenosine opposite the lesion and produced single nucleotide deletion products.[234] Recently, Minko *et al* showed that the α-anomer of

nitrogen mustard-FAPy-dG adducts caused targeted G → T transversion mutations in primate COS7 cellsgeted mutations at a frequency.[238]

Tudek *et al* observed that M13mp18 phage DNA containing methyl-FAPy-dG did not lead to an increase in GC mutations in *E. Coli* following SOS induction.[227] Conversely, methyl-FAPy-dA lesions in the same vector lead to an increase in A → G transition mutations two-fold greater than GC mutation.[227, 253] These results suggested that methyl-FAPy-dA lesions mispairs with 2'-deoxycytidine in SOS-induced *E. Coli* while methyl-FAPy-dG adducts were primarily a lethal lesion.[227] Considering the structural similarity between EB-FAPy-dG and methyl-FAPy-dG, it is conceivable that the formation and accumulation of EB-FAPy-dG lesions could be lethal or lead to replication-induced mutations that will contribute to the observed mutagenicity and toxicity after BD exposure.

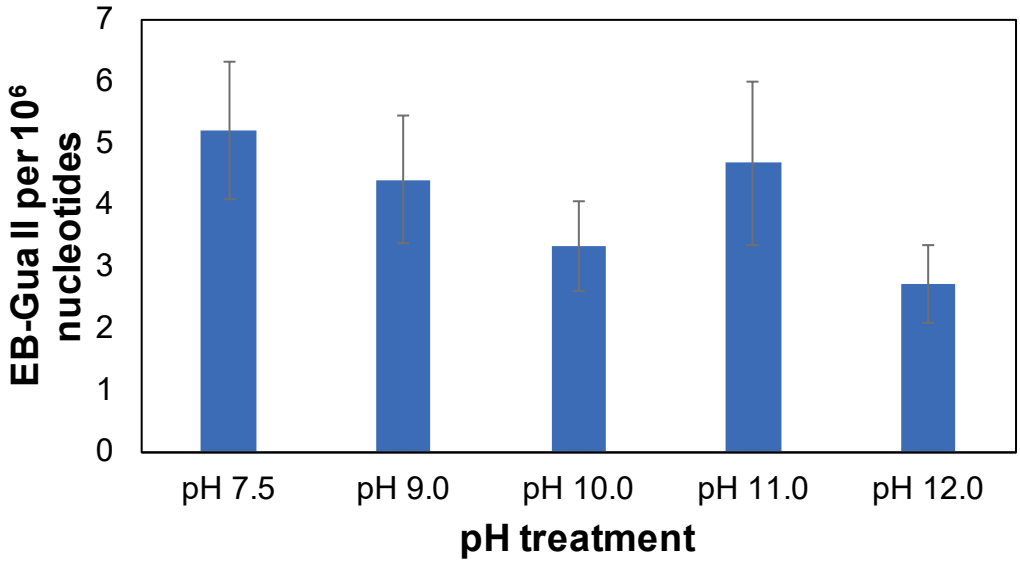
Our cytotoxicity experiments showed NEIL1$^{-/-}$ MEF cells were 1.5-fold more sensitive to EB than NEIL1$^{+/+}$ MEF cells, and therefore EB-FAPy-dG could potentially be a substrate for NEIL1 enzyme. NEIL1 is a BER bifunctional DNA glycosylase enzyme with the known substrates of thymine glycol, FAPy-G, FAPy-A, 8-oxoguanine, 5-hydroxyuracil, dihydroxyuracil, spiroaminohydantioins, guanidinohydantoins, and AFB_1-FAPy-dG.[259, 260, 573] To investigate whether EB-FAPy-dG was a substrate for NEIL1, we treated wild type (NEIL1$^{+/+}$) and NEIL1$^{-/-}$ isogenic control MEF cells with 500 μM for 24 h and quantified EB-FAPy-dG and EB-Gua II adducts. NEIL1$^{-/-}$ cells had 2-fold greater EB-FAPy-dG formation (6.7 ± 0.5 and 3.4 ± 0.03 EB-FAPy-dG per 10^6 nucleotides) and 3-fold greater EB-Gua II levels (14.0 ± 1.5 and 5.3 ± 0.2 EB-Gua II per 10^6 nucleotides) compared to wild type cells (**Figure 5-11**). The increase in both EB-FAPy-dG and EB-Gua

II adduct in NEIL1$^{-/-}$ cells suggests that the NEIL1 enzyme can excise a portion of the formed EB-FAPy-dG within 24 hours.

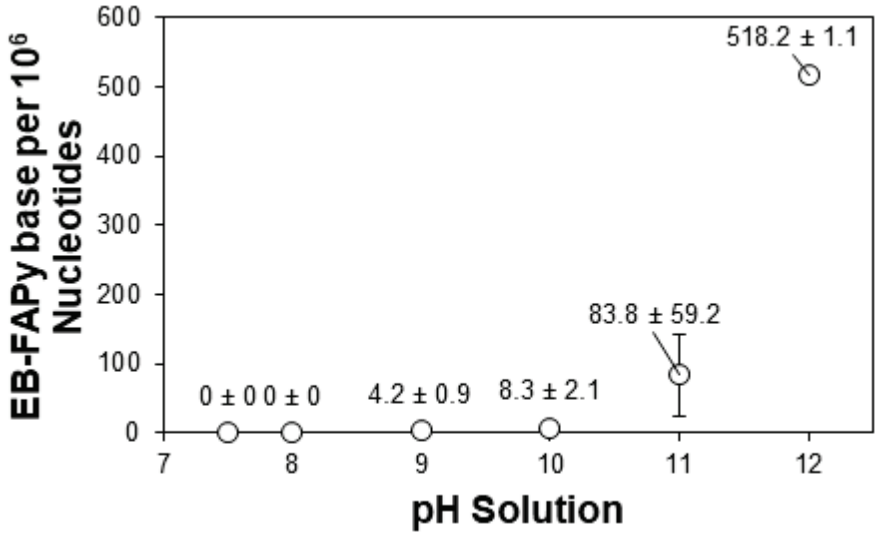
The modest difference in EB-FAPy-dG levels between NEIL1$^{+/+}$ and NEIL1$^{-/-}$ cells suggests EB-FAPy-dG can be excised by additional BER enzyme(s). Structurally analogous alkyl-FAPy-dG adducts have been identified as substrates for other BER enzymes. For example, methyl-FAPy-dG has been identified as a substrate for both formamidopyrimidine DNA glycosylase (FPG) and *E. Coli* endonuclease IV.[254, 569] We observed the selective release of EB-FAPy base from the backbone of EB-treated CT DNA by FPG, which was equally efficient as neutral thermal hydrolysis by 70 ºC for 40 min (unpublished data). However, ethyl-FAPy-dG adducts were found to be excised 7-times slower by endonuclease IV than methyl-FAPy-dG, making it unlikely the much larger EB-FAPy-dG is a substrate.[256] Furthermore, endonuclease IV has been identified as selective for the α-anomer of methyl-FAPy-dG, suggesting it plays a specialized role in FAPy repair in general.[569] Therefore, additional experiments are needed to identify all the enzymes and repair pathway(s) involved in the removal of EB-FAPy-dG.

Figure 5-13: Analysis of EB-Gua II and EB-FAPy-dG formation during different steps of sample processing. (**A**) *In vitro* pH-dependent EB-Gua II formation in CT DNA that was treated with 5mM EB, then incubated at 37 ºC for 3 days to release alkylated dG. (**B**) pH-dependent formation of EB-FAPy base without incubation at 37 ºC for 3 days. No EB-FAPy base was detected at pH 7.5 and 9.0.

A.



B.



5.5. Conclusions

In conclusion, we studied the formation of ring open FAPy adducts in DNA treated with carcinogenic metabolite of 1,3-butadiene, 3,4-epoxy-1-butene. An accurate quantitative nanoLC/ESI$^+$-MS/MS isotope dilution methodology was developed for the detection and quantitation of EB-FAPy-dG in vitro and in vivo. The applicability of this method was demonstrated by quantitation of EB-Fapy adducts in EB treated CT DNA followed by base treatment. We showed that EB-Fapy adducts preferentially form at pH 11.0 and above, but adduct formation does occur at physiologically relevant pH. Finally, we observed a modest increase in EB-FAPy-dG formation in NEIL1$^{-/-}$ MEF cells compared to wild type MEF cells treated with 500 μM EB, suggesting that EB-FAPy-dG is a substrate of the BER enzyme NEIL1. We conclude that the EB-Fapy-dG is an adduct that can form under physiological conditions, and can therefore contribute to the observed genotoxicity and mutagenicity of BD.

6. Summary and Conclusions

DNA is constantly under the threat of chemical modification by various endogenous and exogenous agents, leading to structural modification of nucleobases to form DNA adducts.[3, 78] The formation of DNA adducts threatens the faithful replication of genetic material, as they can be replicated in an error-prone manner or inhibit DNA replication, ultimately leading to mutations or apoptosis, respectively.[1, 2] To restore DNA to its native state, DNA adducts can be recognized and removed by DNA repair enzymes, preventing mutagenesis (**Figure 1-1**).[5, 6] Elucidating the structure and repair mechanisms of each DNA adduct, as well as their effects on DNA replication , is critical in order to fully understand their biological consequences.

DNA-protein cross-links (DPCs) are ubiquitous, super-bulky DNA lesions that form when proteins become irreversibly trapped on chromosomal DNA.[262, 263] DPCs can be induced after exposure to various physical and chemical agents including ionizing radiation,[274, 275] UV light,[276-278] transition metal ions,[279-282] halogenated hydrocarbons,[283, 284] environmental carcinogens,[285] and common anticancer drugs (**Figure 1-11**).[286-289] Additionally, endogenous DPCs can also form as part of epigenetic regulation[589] or as a result of DNA binding proteins (e.g. topoisomerases[290, 291] and DNA polymerases)[29, 30] becoming trapped on DNA intermediates. DPCs can be induced by reactive oxygen species (ROS) produced during cellular respiration and inflammation,[292-294] and by aldehyde byproducts of lipid peroxidation such as methylglyoxal[295] and malondialdehyde.[296] Previous *in vitro* investigations have shown that DPCs interfere with key biological processes including DNA replication[264-267] and transcription.[268] Due to these properties, the

formation and accumulation of DPCs have been proposed to play a role in aging,[269, 270] cancer,[271, 272] heart disease,[269, 270] and neurodegenerative disorders.[273]

Despite recent interest in the formation and biological consequences of DPCs, there are still many unanswered questions. The identity of the protein constituents susceptible to DPC formation and the cross-linking mechanisms have not been fully investigated for many DPC-inducing agents. Furthermore, the extent of DPC formation after exposure to DPC-inducing agents has not been accurately quantified. Finally, the mechanism(s) of active DPC repair remains controversial and is still not fully understood (**Section 1-7**). Towards these goals, in the first part of this Thesis, we investigated the identity of the protein constituents, the atomic connectivity, and the repair mechanism(s) of DPCs induced by *bis*-electrophiles and ROS.

In Chapter 2 of this thesis, we istudied the formation and repair of DPCs induced by the active metabolites of cyclophosphamide, phosphoramide mustard (PM) and nornitrogen mustard (NOR, **Chapter 2**). Cyclophosphamide (CP) is a nitrogen mustard drug used as first-line chemotherapy treatment for cancers including lymphoma,[373, 374] leukemia,[375] and myeloma.[376, 377] CP is bioactivated by the cytochrome P450 monooxygenases CYP2B5 and CYP3A4 to yield 4-hydroxy cyclophosphamide (4-OH-CPA), which spontaneously breaks down to DNA-reactive metabolites PM and NOR, which are both capable of inducing DPCs (**Figure 1-11, Scheme 1-6**).[381, 382]

We first utilized mass spectrometry-based proteomics technique to characterize the proteins susceptible to PM-induced DPC formation in cells (**Section 2.2.4**). Human fibrosarcoma cells (HT1080) were treated with 100 μM PM, and the proteins covalently trapped to chromosomal DNA were isolated using a modified phenol/chloroform

extraction methodology developed by our group. After mass spectrometry based proteomics analysis, a total of 134 proteins were found to form DPCs in the presence of PM (**Table 2-1**). Gene ontology analysis revealed that these proteins are involved in a diverse set of biological processes, including transcriptional regulation, DNA repair, chromosomal organization, apoptosis, and apoptosis (**Figure 2-4**). When our list of PM-induced DPCs was compared to analogous lists of proteins crosslinked to DNA by DEB and cisplatin, a total of 47 proteins (30.9%) and 106 proteins (41.4%) proteins were found in common respectively (**Figure 2-9**). This comparison suggests that nitrogen mustards and platinum drugs may target similar proteins for cross-linking to DNA, while DEB targets a distinct group of protein targets.

HPLC-ESI$^+$-MS/MS analysis of complete proteolytic digests of PM-induced DPCs revealed *N*-[2-[cysteinyl]ethyl]-*N*-[2-(guan-7-yl)ethyl]amine (Cys-NOR-N7G) conjugates, confirming that PM-induced cross-linking occurred between the cysteine thiols of proteins and the N7-position of guanines in DNA. Authentic standards and internal standards ($^{15}N_5$-labeled guanine) of Cys-NOR-N7G were synthesized and used to develop a sensitive isotope dilution tandem mass spectrometry assay to perform absolute quantitation of PM-induced DPC formation (**Section 2.2.1** and **Section 2.2.5**). When HT1080 cells were treated with an increasing amount of PM (0 – 1000 μM), a dose-dependent increase in Cys-NOR-N7G was observed, confirming the assay could be utilized for quantifying PM-induced DPC formation (**Figure 2-7**).

To begin elucidating the repair mechanism of PM-induced DPCs, we conducted cytotoxicity experiments using human cell lines deficient in a single DNA repair pathway and their corresponding isogenic controls (**Section 2.2.2**). Both human NER-deficient

(XPA and XPD cells) and HR-deficient (FANCD2 cells) were twice as sensitive to PM compared to wild type controls, suggesting these pathways may be involved in the repair of PM-induced DPCs (**Table 2-2**). To determine whether the observed sensitivity differences were caused by impaired removal of PM-induced DPCs, the cell lines were treated with 250 μM NOR, then allowed to recover for 4 hours to allow DPC repair. Quantitation of Cys-NOR-N7G revealed that the NER-deficient cell lines had significantly greater DPC levels (24.3 ± 3.5 and 18.7 ± 4.3 adducts/10^8 nucleotides for XPA and XPD cell lines respectively) as compared to XPD corrected isogenic controls (10 ± 2.3 adducts/10^8 nucleotides), HT1080 (8.7 ± 1.5 adducts/10^8 nucleotides), PD20 (5.2 ± 1.3 adducts/10^8 nucleotides), and PD20-corrected cells (5.6 ± 1.6 adducts/10^8 nucleotides) (**Figure 2-8A**). Furthermore, we measured changes in NOR-induced DPC repair over time by treating HT1080, XPA, PD20, and PD-corrected cells with NOR (250 μM), followed by allowing the cells to recover in drug-free media for 0, 4, 12, or 24 h. HPLC-ESI^+-MS/MS analysis revealed that XPA cells had the highest Cys-NOR-N7G adduct levels at each time point measured (14.7 ± 6.2, 19.2 ± 5.1, 10.2 ± 6.1, and 7.2 ± 3.3 at 0, 4, 12, and 24 h, respectively (**Figure 2-8B**). Taken together, these results are consistent with nucleotide excision repair being required for active repair of cyclophosphamide-mediated DPCs.

Reactive oxygen species (ROS) are radical and non-radical oxygen species formed by a partial reduction of oxygen. Examples of ROS include superoxide anions (O_2^-), hydrogen peroxide (H_2O_2), and hydroxyl radicals ($^{\bullet}OH$). ROS can be produced endogenously in cells during aerobic cellular respiration,[338-340] immune response,[341, 342] and inflammation,[35] and are used as intracellular messengers in cellular signaling pathways.[349-

[355] However, ROS can also oxidatively damage cellular biomolecules including DNA and proteins, leading to the direct covalent attachment of proteins to DNA to form DPCs. Some oxidative DNA lesions can undergo secondary reactions with nucleophilic sites of proteins, leading to the formation of hydrolytically stable DPCs.[263, 356] Alternatively, DPCs can form directly via one electron addition reactions after the oxidation of DNA bases. For example, one electron oxidation of guanine can form radicals that react with lysine side chain of proteins to yield N^{ε}-(guanin-8-yl)-lysine cross-links.[361] Similarly, hydrogen abstraction from the C5-methyl group of thymidine yields 5-(uracilyl)methyl radical, which can undergo a radical addition reaction to the C-3 position of a neighboring tyrosine molecule to form thymidine-tyrosine cross-links.[362, 363]

One biological event that stimulates a significant influx of ROS in the heart is myocardial infarction/reperfusion (MI, heart attack). MI occurs when the blood flow is decreased or stopped in a cardiac artery, restricting the delivery of oxygenated blood to the affected myocardium. Current MI therapies are based on reperfusion, or the re-opening, of the blocked artery, returning blood flow to the damaged myocardium. However, the sudden influx of oxygenated blood is associated with a rapid influx of ROS to the already damaged myocardium.[366-368] During the MI, the lack of oxygen switches cardiomyocytes and cardiac fibroblast cells to anaerobic respiration mode.[367, 369] Reperfusion-induced oxygen return triggers the sudden reactivation of the electron transport chain, producing a spike in ROS that can permeate from the mitochondrial matrix.[367, 369] Furthermore, neutrophils accumulating at the damaged myocardium mediate tissue damage by releasing matrix-degrading enzymes that release ROS.[370-372] Despite the evidence that MI/reperfusion

significantly increase ROS levels in myocardium, at the time of this investigation there was no evidence that this leads to an increase in ROS-induced DPC formation.

In Chapter 3 of this thesis, we investigated ROS-induced DPC formation caused by MI/reperfusion using the model of left anterior descending artery ligation/reperfusion surgery in rat.[468, 487] To serve as controls, additional rats were subjected to a sham surgery. Immediately after reperfusion, cardiomyocytes and cardiac fibroblasts were separated by Langendorff reperfusion,[590] allowing us to focus on DPC formation in cardiomyocytes, the functional unit of the heart. Rats that underwent LAD ligation/reperfusion exhibited elevated plasma levels of cTnI (>50 ng/mL), confirming severe MI (**Table 3-1**). Furthermore, plasma levels of oxidative stress biomarker 8-iso-PGF2α were increased 3-5-fold in animals that underwent the LAD surgery as compared to sham surgery (**Figure 3-1**), confirming that the reperfusion step led to an influx of ROS. Taken together, these results confirmed that our LAD ligation/reperfusion animals experienced an MI/reperfusion injury and could be further investigated.

Previous research has demonstrated that the total digestion of ROS-induced DPCs yields thymine-tyrosine conjugates, 2-amino-3-(4-hydroxy-3-((1-((2R,4R,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)methyl)phenyl)propanoic acid (dT-Tyr). We synthesized authentic dT-Tyr and [$^{15}N_1$,$^{13}C_9$]- dT-Tyr that were used to develop a sensitive and accurate isotope dilution nanoLC-ESI$^+$-MS/MS assay to perform absolute quantitation of ROS-induced DPCs. We found that dT-Tyr amounts were significantly increased in cardiomyocytes of rats that underwent LAD ligation/reperfusion as compared to the sham group (35.7 ± 2.2 vs 18.2 ± 1.5 dT-Tyr adducts per 10^8 nucleotides respectively, **Figure 3-2B**), supporting our

hypothesis that ischemia-reperfusion injury increases the levels of ROS-mediated DPCs in cardiomyocyte DNA.

To simultaneously identify the proteins susceptible to DPC formation and quantify the changes in ROS-induced DPC formation, tryptic peptides were labeled with Themo Mass Tags (TMT) and performed quantitative proteomics. A total of 90 proteins participating in ROS-induced DPC formation were characterized, including ROS scavengers, contractile proteins, and regulators of apoptosis (**Table 3-2**). The majority of these proteins (N = 80, 88.9%) were increased in abundance at least 1.2-fold after LAD ligation/reperfusion as compared to sham surgery, with 21 (23.3%) of these proteins increasing in abundance at least 5-fold (**Table 3-2**). Quantitative proteomics analysis of the global proteome changes was significantly less pronounced, confirming that the increase in DPC formation after LAD ligation/reperfusion were not due to an overall increase in protein expression but rather an increase in DNA-bound proteins due to ROS production.

Although significant progress has been made recently toward elucidating the cellular mechanisms of DPC repair, many questions remain unanswered. Specifically, Spartan metalloprotease was proposed to play a key role in replication-coupled repair of DPCs.[428, 431, 512] Spartan is a mammalian homolog to the yeast Wss1, as Spartan shares both similar domain organization and a common evolutionary origin.[428] Furthermore, purified Spartan protein has been shown to proteolytically digest the DNA-bound proteins histone H1, H2A, H2B, H3, and Hmg1 in the presence of ssDNA.[431] However, at the time this investigation started, there was no direct evidence confirming that Spartan in involved in repair of chromosomal DPCs *in vivo*.

In Chapter 4 of this thesis, we utilized our nanoLC-ESI^{+}-MS/MS assay for dT-Tyr to quantify ROS-induced DPC formation in tissues of 10-month-old wild type (*SPRTN* $^{+/+}$) and hypomorphic (*SPRTN* $^{H/H}$) mice. A significant (1.5 – 2-fold) increase in ROS-induced DPCs was observed in the brain, liver, heart, and kidney tissues of Spartan deficient mice (*SPRTN* $^{H/H}$) as compared to their wild type controls. To our knowledge, this is the first direct evidence that Spartan is involved in repair of ROS-induced DPCs in normal tissues. Interestingly, we were unable to detect dT-Tyr in the spleen or thymus tissues, suggesting that in these tissues, there is an additional mechanism for DPC processing, e.g. immunoproteasome,[545] this hypothesis required additional experimnets to be tested.

In the second section of this Thesis (Chapter 5), we elucidated the structures of novel DNA adducts induced by human carcinogen 1,3-butadiene (BD). BD is a colorless, volatile gas commonly present in automobile emissions, urban and industrial air, and cigarette smoke.[172, 177-179] BD is also an important raw material used in the production of synthetic rubber and plastics, especially tires used in the automobile industry.[178, 179] Epidemiological studies have found that exposure to BD is linked to an increased risk of leukemia, lymphatic and hematopoietic cancers, cardiovascular disease, and respiratory disorders.[183-190]

BD is metabolically activated by cytochrome P450 monooxygenases to three inherently reactive epoxides: 3,4-epoxy-1-butene (EB), 1,2,3,4-diepoxybutane (DEB), and 1,2-dihydroxy-3,4-epoxybutane (EBD).[192] These electrophilic species can react with nucleophilic bases of DNA to form covalent nucleobase adducts,[558] which are thought to be responsible for BD-induced mutagenesis and carcinogenicity.[196, 197, 559-562] Previously identified and structurally characterized BD-DNA monoadducts include N-7-(2-hydroxy-

3-buten-1-yl)-guanine (EB-Gua I),[200, 563] N-7-(1-hydroxy-3-buten-2-yl) guanine (EB-Gua II),[200, 563] N7-(2,3,4-trihydroxybut-1-yl)-guanine (N7-THBG),[206, 207] N1-(2,3,4-trihydroxybut-1-yl)-adenine (N1-THB-Ade),[564] 1,4-*bis*-(guan-7-yl)-2,3-butanediol (*bis*-N7G-BD)[565,566] and 1,N^6-(1-hydroxymethyl-2-hydroxypropan-1,3-diyl)-2′-deoxyadenosine (1,N^6-HMHP-dA).[209, 567]

Alkylation of the N7-position of guanine by EB initially generates the positively charged DNA adduct N7-EB-dG.[77, 252] This adduct is hydrolytically labile because of the intrinsic destabilization of the glycosidic bond, and is therefore subject to spontaneous depurination to release EB-Gua from the DNA backbone.[33]Alternatively, the C8 position of N7-EB-dG can be attacked by a hydroxyl anion, leading to imidazole ring opening and the formation of N^6-(2-Deoxy-D-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-N-(2-hydroxybut-3-en-1-yl)formamidopyrimidine (EB-Fapy-dG) adducts (**Figure 1-9**).[233] However, at the time of this investigation, no BD-induced FAPy-dG adducts had been identified or structurally characterized. Furthermore, it was unknown whether these adducts contribute to the observed mutagenic and genotoxic effects of BD.

In Chapter 5 of this thesis, we synthesized authentic standards and internal standards of N^6-(2-Deoxy-D-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-*N*-(2-hydroxybut-3-en-1-yl)formamidopyrimidine (EB-Fapy-dG), and structurally characterized these standards by UV-vis, ESI^+-MS/MS, ^{1}H-NMR, and ^{1}H-^{13}C HSQC (**Figures 5-2A, 5-2B, 5-3A, and 5-3B**). Our NMR analysis revealed that EB-FAPy-dG could exist as multiple isomers in solution, including atropisomers caused by restricted rotation around the C5-N^5 bond and distinct rotamers caused by rotation around the formyl N-CO bond. However, our NMR analysis revealed that EB-FAPy-dG does not

spontaneously interconvert between β- and α-anomers as observed with other FAPy-dG adducts.[224, 587]

The synthetic EB-FAPy-dG and its ^{15}N-labeled analogue were used to develop a nanoLC-ESI$^+$-MS/MS assay in order to detect EB-FAPy-dG formation in EB-treated DNA and living cells. Using this novel assay, we observed a linear increase in EB-FAPy-dG adducts in CT DNA treated with increasing concentrations of EB under basic conditions (**Figure 5-5**). Furthermore, we were able to detect *in vitro* EB-FAPy-dG formation at pH 7.5 (**Figure 5-6**), proving that EB-FAPy-dG could form under physiologically relevant conditions. However, the highest EB-FAPy-dG adduct levels were detected at pH 12.0 (**Figure 5-6**).

Recently, the human base excision repair (BER) enzyme endonuclease 8-like protein 1 (NEIL1) was identified as the enzyme that excises the bulky aflatoxin B_1-FAPy-dG adduct.[573] Given that NEIL1 has also been implicated as the BER enzyme that removes Me-FAPy-dG in mammalian cells,[572] we hypothesized that EB-FAPy-dG could be a substrate for NEIL1 mediated repair. When NEIL1$^{-/-}$ and isogenic corrected (NEIL1$^{+/+}$) MEF cells were treated with 500 μM EB for 24 hours, we detected increased formation of EB-FAPy-dG in the NEIL1-deficient cells (4.45 ± 0.51 EB-FAPy-dG per 10^6 nucleotides) as compared to repair proficient wild type fibroblasts (2.27 ± 0.032 ± 0.03 EB-FAPy-dG per 10^6 nucleotides, **Figure 5-8**).

In summary, the studies presented herein employed a combination of organic synthesis and mass spectrometry based analytical methodologies to investigate the formation and repair of DNA lesions *in vitro* and *in vivo*. Mass spectrometry-based proteomics techniques were utilized to identify the proteins susceptible to PM- and ROS-

induced DPC formation (**Chapters 2** and **3**). The atomic connectivity of these lesions was then confirmed by analyzing the complete digestion products of PM- and ROS-induced DPCs (Cys-NOR-N7G and dT-Tyr respectively) by HPLC-ESI$^+$-MS/MS (**Chapters 2** and **3**). Authentic standards and internal standards of Cys-NOR-N7G and dT-Tyr were synthesized and to develop sensitive and accurate isotope dilution tandem mass spectrometry assays to perform absolute quantitation of PM- and ROS-induced DPCs (**Chapters 2 – 4**). Specifically, these assays were used to confirm MI/reperfusion leads to an increase in ROS-induced DPCs (**Chapter 3**) and to begin elucidating the repair mechanisms of PM- and ROS-induced DPCs. Overall, our initial repair experiments suggest that NER (**Chapter 2**) and the metalloprotease Spartan (**Chapter 4**) are crucial for DPC repair in vivo. Finally, we successfully synthesized and structurally characterized a novel BD-induced DNA adduct EB-FAPy-dG (**Chapter 5**) and developed a sensitive isotope dilution tandem mass spectrometry assay for its detection in cells and treated DNA (**Chapter 5**). To our knowledge, this is the first report of a BD-induced FAPy adduct, and future studies will examine whether BD-induced FAPy adducts contribute to mutagenicity and carcinogenicity of BD.

7. Future Directions

7.1. Investigating the cardioprotection potential of the natural products Ruscogenin and Ophiopogonin D via preventing ROS-induced DPC formation after reperfusion injury

As described in **Chapters 3** and **4**, we have developed a sensitive stable isotope dilution nanoLC-MS/MS assay for the absolute quantification of endogenous ROS-induced and IR-induced DNA-protein cross-linking in cells and tissues. ROS- and IR-induced DPCs are formed by hydroxyl-radical hydrogen abstraction from the 5-methyl position of a thymidine base to yield a reactive thymidine radical, which undergoes a one-electron addition to the 3-position of a tyrosine to yield a stable methylene linkage (**Scheme 3-2**). Subsequent hydrogen abstraction from the 3-position of tyrosine re-aromatizes the phenol ring to yield a stable DNA-protein cross-link (**Scheme 3.2**). The protein and DNA constituents of IR- or ROS-induced DPCs can be enzymatically digested to 2-amino-3-(4-hydroxy-3-((1-((2R,4R,5R)-4-hydroxy-5-(hydroxy-methyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)methyl) phenyl)propanoic acid (dT-Tyr, **Scheme 3-1**), which can serve as a biomarker of ROS-induced DPC formation.

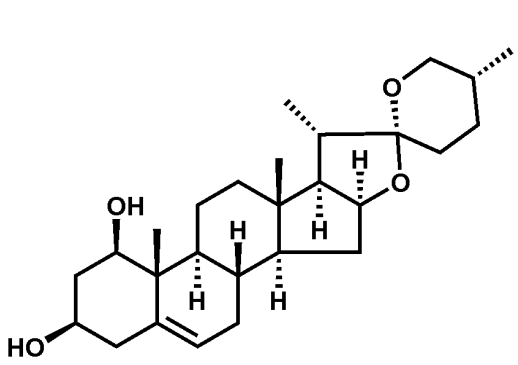
Using our nanoLC-ESI$^+$-MS/MS dT-Tyr assay, we observed an increase in ROS-induced DPC formation in cardiomyocytes following LAD ligation/reperfusion (**Figure 3-2**). The assay was sensitive enough to detect changes in ROS-induced DPC formation caused by reperfusion injury. Therefore, the same methodology could be used to investigate the efficacy of novel antioxidant therapies to prevent reperfusion-induced ROS accumulation and subsequent DPC formation.

Extracts of the plant *Ophiopogon japonicus* have been used for centuries in traditional Chinese herbal medicine to treat cardiovascular disease. Studies have shown that *O. japonicus* extract may provide cardiovascular protection through anti-arrhythmia,[591] ROS scavenging,[592, 593] inhibiting platelet aggregation,[594] and improving microcirculation.[594] One active component of *O. japonicus*, the steroidal glycoside Ophiopogonin D (**Figure 7-1**), has been shown to partially protect endothelial cells from H_2O_2-induced oxidative stress[595] and to prevent doxorubicin-induced cardiac toxicity.[596] Our collaborator Professor Boyang Yu (Chinese Pharmaceutical University) is currently investigating an additional steroidal glycoside, Ruscogenin (**Figure 7-1**), which has anti-inflammatory and anti-arrhythemia properties.[594, 597] However, neither Ophiopogonin D nor Ruscogenin have been previously investigated as an antioxidant therapy to mitigate reperfusion injury.

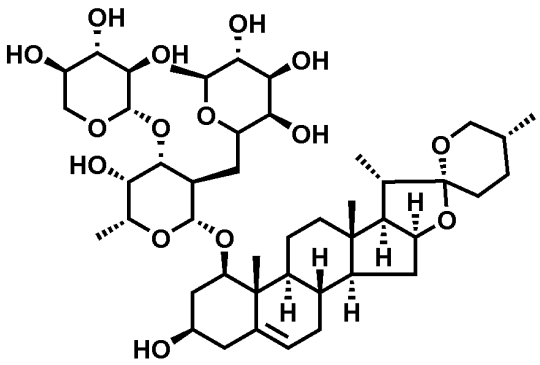
We propose to investigate Ophiopogonin D and Ruscogenin's potential as a preventative antioxidant therapy following ischemia/reperfusion injury. A set of age-matched Sprague Dawley rats (N = 5) will be administered either Ophiopogonin D or Ruscogenin (5, 10, or 20 mg/kg) via intraperitoneal injection every other day for seven consecutive days.[596] The other set of rats (N = 5) will remain untreated. Following the pretreatment period, all rats will undergo an LAD ligation/reperfusion procedure, followed by the isolation of cardiomyocytes by Langendorff reperfusion and extraction of ROS-induced DPCs as described in **Section 3.2**. ROS-induced DPCs will then be enzymatically digested, and dT-Tyr will be quantified as a biomarker of ROS-induced DPC formation (**Section 3.2**). If pretreatment with the two natural products successfully reduces ROS formation, we would expect to see a decrease in dT-Tyr levels after pretreatment compared

to LAD ligation/reperfusion only at every concentration investigated. If pretreatment with the steroidal glycosides does not prevent ROS accumulation, we would expect to observe similar dT-Tyr levels as observed in **Section 3.3.2** regardless of the condition.

Figure 7-1: Structures of the cardioprotective Ruscogenin and Ophiopogonin D, active metabolites found in *O. japanicus* extract.



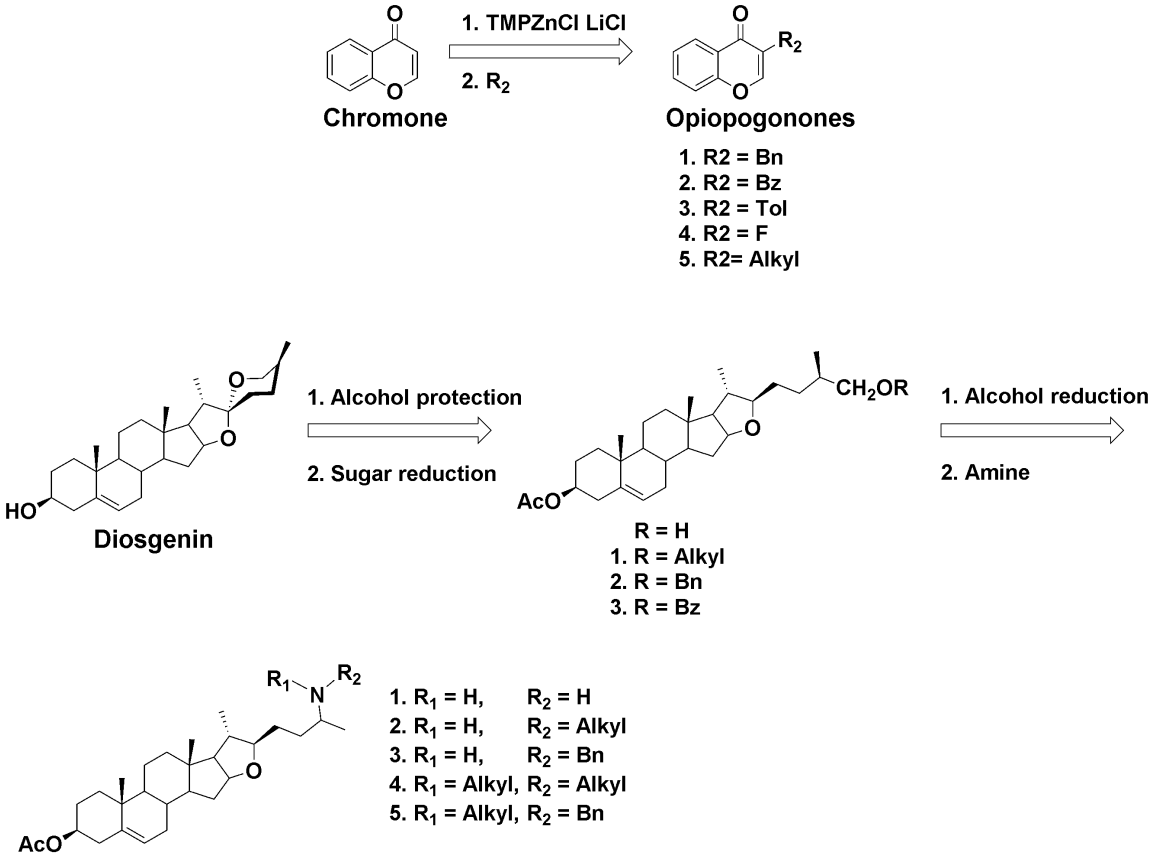
Ruscogenin



Ophiopogonin D

If either Ophiopogonin D or Ruscogenin successfully prevents ROS-induced DPC formation, a series of structural analogs can be analyzed to perform a structure-activity relationship (SAR) study. For example, a series of novel Ophiogponones, a structurally similar active metabolite found in *O. japonicus* extract that shares a 3-benzyl chromone skeleton,[598] can be synthesized from commercially available chromone via regioselective zincation (**Scheme 7-1**).[599, 600] Similarly, a series of Furostane derivatives, another active metabolite in *O. japonicus* extract that shares a steroidal backbone,[598] can be synthesized by opening the spiroketal moiety of commercially avaialable Diosgenin.[601] The efficacy of each novel compound can be first tested by pretreating cardiomyocytes in culture with the novel glycoside, followed by treating cells with hydrogen peroxide and measuring changes in ROS-induced DPC formation by dT-Tyr quantitation. As a control, dT-Tyr levels will be quantified in cells pretreated with Ophiopogonin D or Ruscogenin and treated with hydrogen peroxide. Only analogs that decrease DPC formation better than the controls will be further investigated in the animal model experiments described above.

Scheme 7-1: Synthetic strategy for synthesizing series of Ophiopogonin D and Ruscogenin analogs.



7.2. Quantitative proteomic analysis and comparison between chromosomal DPCs and mitochondrial DPCs induced by reperfusion injury

As described in **Section 3.3.4**, we utilized quantitative proteomics to simultaneously identify and relatively quantify ROS-induced DPC formation after reperfusion injury in a rat model of myocardial infarction/reperfusion (LAD ligation/reperfusion). A total of 90 proteins were found to participate in ROS-induced DPC formation, with increased crosslinking following the simulated heart attack. Similar analysis of global proteome abundance revealed very modest changes in protein expression after LAD ligation/reperfusion compared to sham surgery, confirming that the increases in DPC formation after LAD ligation/reperfusion was caused by an increase in ROS-mediated cross-linking and not total protein abundance (**Table 3-3**). However, our DPC extraction methodology completely lyses both the nuclei and the mitochondrial membranes of cardiomyocytes, and therefore cannot discriminate mitochondrial DPCs (mtDPCs) from chromosomal DPCs (cDPCs). Considering that DPCs are known to interfere with DNA replication[264, 266, 267] and transcription,[268] and the formation and accumulation of DPCs are thought to contribute to the etiology of cardiovascular disease after myocardial infarction,[170, 171, 270] it is important to distinguish between mitochondrial and chromosomal DPCs.

To identify proteins only covalently cross-linked to either chromosomal DNA or mitochondrial DNA, we can modify our DPC extraction methodology to separate nuclei and mitochondria from other organelles. In short, cardiomyocytes can be lysed using a weak hypotonic cell lysis solution that breaks the cell membrane but maintains the integrity

of nuclei and mitochondria.[288, 440] Organelles can then be separated by ultracentrifugation (> 20,000 RPM) through sucrose gradients following established protocols.[602, 603] With the nuclei and mitochondria isolated from each other, cDPCs and mtDPCs can be separately extracted using Qiagen cell lysis solution and the protocol described in **Section 3.2**. Tryptic digestion and TMT labeling of the cDPCs and mtDPCs would allow for simultaneous identification of the proteins susceptible to ROS-induced DPC formation in both organelles, and allow for relative quantitation to identify the proteins most susceptible to DPC formation (**Section 3.2**). Relative increases in ROS-induced DPC formation in both organelles can also be confirmed by absolute quantitation of dT-Tyr using our nanoLC-ESI^+-MS/MS methodology described above.

We hypothesize that many of the mitochondrial proteins found in **Table 3-2** will be identified as constituents of cDPCs. As described in **Section 3.4**, several proteins such as pyruvate dehydrogenase are known to be recruited to the nucleus during reperfusion injury generate nuclear acetyl CoA, which is required for histone acetylation and the activation of transcription. In contrast, other proteins such as Enoyl-CoA hydratase, 2,4-dienoyl-CoA reductase, and ATP synthase have no known function in the nucleus. To confirm the translocation of these mitochondrial proteins under the conditions of oxidative stress, future fluorescence microscopy experiments can be conducted. Specifically, primary cardiomyocytes can be treated with hydrogen peroxide (10 mM) and cells will be fixed at various times post treatment for primary antibody incubation. Incubation with a FITC-labeled secondary antibody would then allow visualization of the protein by standard fluorescence microscopy technique.[604] Co-staining with 4',6-diamidino-2-phenylindole (DAPI) and Rhodamine 123 would confirm the location of the nucleus and mitochondria,

respectively, and confirm the translocation of the proteins of interest from the mitochondria to the nucleus. We anticipate that mitochondrial proteins identified as constituents of cDPCs will be localized to mitochondria under the conditions of oxidative stress.

7.3. Elucidate the role of Spartan metalloprotease in the repair of ROS-induced DPCs in cardiomyocytes

As discussed in **Section 1.7** and **Chapter 4**, Spartan metalloprotease is currently being investigated as the enzyme that digests the protein constituent of a DPC to smaller lesions to allow for their repair. Our collaborator Prof. Yuichi Machida (Mayo Clinic) has detected Spartan protein in the liver, brain, lung, and kidney of mice by western blot analysis,[514] but did not investigate Spartan expression in the myocardium. In **Section 4.3.4**, we observed an increase in dT-Tyr in the myocardium of $SPRTN^{H/H}$ mice compared to wild type mice (46.02 ± 6.08 and 29.89 ± 3.68 adducts per 10^8 nucleotides respectively, see **Figure 3-2B**), suggesting that Spartan may play a role in the repair of ROS-induced DPCs in cardiomyocytes and cardiac fibroblasts.

To test Spartan's presence in myocardium, we will first isolate cardiomyocytes and cardiac fibroblasts from the hearts of wild type mice as described in **Section 3.2**, followed by standard western blot analysis of the proteome using mouse monoclonal anti-Spartan antibody (IgG_{2a}, clone 11A).[539] Based on our nanoLC-ESI^+-MS/MS data dT-Tyr conjugates presented in in **Chapter 4** (**Figure 4-6B**), we anticipate to detect Spartan protein in mouse cardiomyocytes and/or cardiac fibroblasts. Additionally, mRNA can be isolated from both cell types using a commercially available kit (Oligotex mRNA mini kit, Qiagen, Hilden Germany), and SPRTN mRNA levels can then be quantified using Real-Time Quantitative Polymerase Chain Reaction (RT-qPCR) to confirm Spartan expression.[605] Spartan expression levels have been previously measured using the forward and reverse primers 5′-GGA CCT TGT AGA GAC TCT TTT G-3′ and 5′-CCT CAT CAT GGA AAG TGT GG-3′ respectively.[514] Given the limited sensitivity of the currently available Spartan

antibody,[429, 514, 539] it is possible that we will be unable to detect Spartan by western blot but will observe SPRTN gene expression.

In the current model of DPC repair, Spartan mediated DPC repair is linked to DNA replication because it was shown to be recruited to stalled replication forks (**Chapter 1.7**).[430, 431, 513] Since cardiomyocytes cannot replicate, a different repair mechanism must be in play in cardiac tissues. Our future studies will investigate possible recruitment of Spartan to DPC lesions in non-replicating cells. Furthermore, depletion of Spartan by siRNA methodology could be performed to investigate Spartan's role in maintaining the genomic stability of myocardium.[429]

7.4. Determine the contribution of the proteasome in the digestion of the protein constituent of DPCs in DPC repair

Using our sensitive dT-Tyr nanoLC-MS/MS assay, we observed a slight increase (1.5 – 2 fold) in dT-Tyr conjugates in the brain, liver, heart, and kidneys of SPRTN-deficient hypomorphic mice compared to wild-type mice (**Figure 4-6B** in **Chapter 4**). Furthermore, we were unable to detect dT-Tyr adduct in the thymus and spleen tissues. Our data suggests that the metalloprotease Spartan is not the only protease involved in the digestion of the protein constituents of DPCs to allow for repair. Although Nakano *et al* was unable to detect ubiquitinated FA-induced DPCs by western blot,[424, 425] studies in other laboratories suggested that the proteasome may play a role in DPC repair when Spartan is compromised.[309, 426, 427] Specifically, In contrast, Baker *et al* and Quievryn *et al* observed a reduced repair of a Dmnt1-containing plasmid[427] and FA- and Cr(IV)-induced DPCs[309, 522] in cells when the 26S proteasome was pharmacologically inhibited. Using our nanoLC-MS/MS assay, we will test whether the pharmacological inhibition of the proteasome in cultured cells will lead to an increase in dT-Tyr adduct levels.

Furthermore, in collaboration with Dr. Deborah Ferrington (UMN, Dept. of Opthamology and Visual Neuroscience), we are also investigating whether mouse knockouts of the 20S proteasome subunit beta-1i, low molecular mass protein-7, or multicatalytic endopeptidase complex subunits (LMP2, LMP7, and MECL genes respectively), which are catalytic subunits of the immunoproteasome, will lead to an increase in dT-Tyr in mouse tissues. Taken together, an increase in dT-Tyr adduct in cells and tissues with reduced proteasome activity will provide direct evidence that the proteasome plays a role in the repair of endogenous ROS-induced DPCs.

To investigate the role of the proteasome in DPC repair, we will pretreat MEF5 and MEF7 cells with lactacystin, a potent small molecule inhibitor of the 20S subunit of the proteasome (**Section 4.3.2)**. The cells will then be irradiated as described in **Section 4.2** and incubated an additional 48 hours to allow the repair of any IR-induced damage. Following 48 hours, IR-induced DPCs will be extracted, and dT-Tyr will be quantified by nanoLC-MS/MS as described in **Chapter 3** and **4**. Control cells will be irradiated to but not treated with inhibitor. We expect to observe a modest increase in dT-Tyr levels in irradiated MEF7 cells compared to MEF5 cells, providing additional evidence that Spartan plays a role in global DPC repair. When the cells are pretreated with lactacystin, we expect to see a large increase in dT-Tyr levels in the irradiated MEF7 cells, but only a modest increase in dT-Tyr MEF5 cells compared to the irradiated controls. This is because MEF7 cells possess only one floxed copy of the SPRTN gene, leading to a decrease in Spartan levels. Therefore, MEF7 cells pretreated with lactacystin will be deficient in two proteases capable of digesting the protein constituent, therefore leading to the largest accumulation of IR-induced DPCs.

As described in **Section 5.3.3**, we were unable to observe dT-Tyr analyte in mouse spleen and thymus tissues, indicating that the immunoproteasome present in lymphatic tissues may participate in DPC repair. To confirm the contribution of the immunoproteasome, we will quantify dT-Tyr levels in the spleen and thymus of wild type and $LMP2^{-/-}$, $LMP7^{-/-}$, and $MECL^{-/-}$ mice. We expect to observe an increase in dT-Tyr levels in both the spleen and thymus of mice deficient in either subunit, indicating that the immunoproteasome contributes to the repair of DPCs. Failure to observe any increase in

dT-Tyr adduct levels in the knockout mice would indicate that another protease(s) must contribute to the digestion of DPCs.

7.5. Identify additional biomarkers of ROS-induced DPC formation

Our studies described in **Chapter 3** and **4** employed thymidine-tyrosine conjugates (dT-Tyr, **Scheme 3-1**) as biomarkers of ROS-induced DPC formation. However, exposure to hydroxyl radicals and other reactive oxygen species induces many other DNA lesions, including electrophilic species capable of DPC formation.[31, 34] As described in **Section 1.2** and **Chapter 4**, cross-links between lysine and spiroiminodihydantoin (5-Lys-Sp and 8-Lys-Sp),[293] lysine and guanidinohydantoin,[294] and serine and oxanines (Ser-Oxa)[606] have been structurally characterized *in vitro* (**Scheme 3-1** respectively). To our knowledge, these types of DPCs have not been detected in living cells or tissues. In our future studies, we can develop sensitive isotope dilution tandem mass spectrometry assays for these additional DPC lesions. Given that spirodihydantoins have been shown to be stable at physiological pH, and lysine is one of the most abundant residues in proteins, 5-Lys-Sp and 8-Lys-SP (**Scheme 3-1**) is anticipated to one of the more abundant ROS-induced DPC cross-links.

Authentic standards of both 5-Lys-Sp and 8-Lys-Sp can be simultaneously synthesized by reacting a mixture of 2'-deoxyguanosine/*N*-acetyllysine with Na_2IrCl_6 and 300 watt IR, while singlet oxygen oxidation of 2'-deoxyguanosine with Rose Bengal or methylene blue in the presence of *N*-acetyllysine and 300 watt IR exclusively produces 5-Lys-Sp.[293] The published high flow LC-MS method for 5-Lys-Sp[293] can serve as our starting point for the development of a more sensitive nanoLC-ESI^+-MS/MS. Both 5-Lys-Sp and 8-Lys-Sp can be detected simultaneously. Furthermore, the methodologies for enzymatic digestion of lesion-containing DNA and enrichment of both lesions has been previously developed using a synthetic DNA oligonucleotide containing the lesion. Due to

the increased sensitivity of nanoLC-ESI$^+$-MS/MS, we anticipate that it can be used to study the formation of these adducts in cells treated with hydrogen peroxide and in animal tissues.

7.6. Derivatization of EB-FAPy-dG to improve sensitivity of nanoLC-ESI$^+$-MS/MS assay

As described in **Chapter 5**, we have developed a sensitive and specific nanoLC-ESI$^+$-MS/MS assay to detect butadiene-induced EB-FAPy-dG adducts *in vitro*. Using this assay, we detected EB-FAPy-dG formation in EB-treated CTDNA at physiological pH (**Section 5.3.3**) and in EB-treated mammalian cells (**Section 5.3.4**). However, the LOQ of EB-FAPy-dG is 25 fmol on column (**Figure 5-6**), showing that this methodology is insufficient for *in vivo* experiments. Furthermore, we were only able to detect EB-FAPy-dG in cells treated with relatively high concentrations of EB (> 100 μM), which is significantly higher than exposure levels expected in animal experiments.

In order to detect EB-FAPy-dG formation in cells treated with physiologically relevant levels of EB, or in animals or individuals exposed to BD, we need to improve the LOQ of the assay to sub-fmol levels. We attributed the low sensitivity of our HPLC-ESI-MS/MS method for EB-FAPy-dG to poor ionization of the molecule under ESI conditions. Changing to different HPLC solvents and columns did not improve the sensitivity (**Section 3.5**). Previous analysis of poorly ionizable analytes such as steroids, glycosides, ribosides, and nucleosides have utilized chemical derivatizations to improve both chromatography and sensitivity.[607-610] However, we failed to successfully derivatize the formamide group of EB-FAPy-dG, and successful protection of the 2'-exocyclic amine with *N,N*-dimethylformamide dimethyl acetal significantly decreased sensitivity (unpublished results).

Considering the 2'-exocyclic amine could be protected on both dG and EB-FAPy-dG, an acetal or aldehyde with a permanent positive charge could be used to derivatize EB-

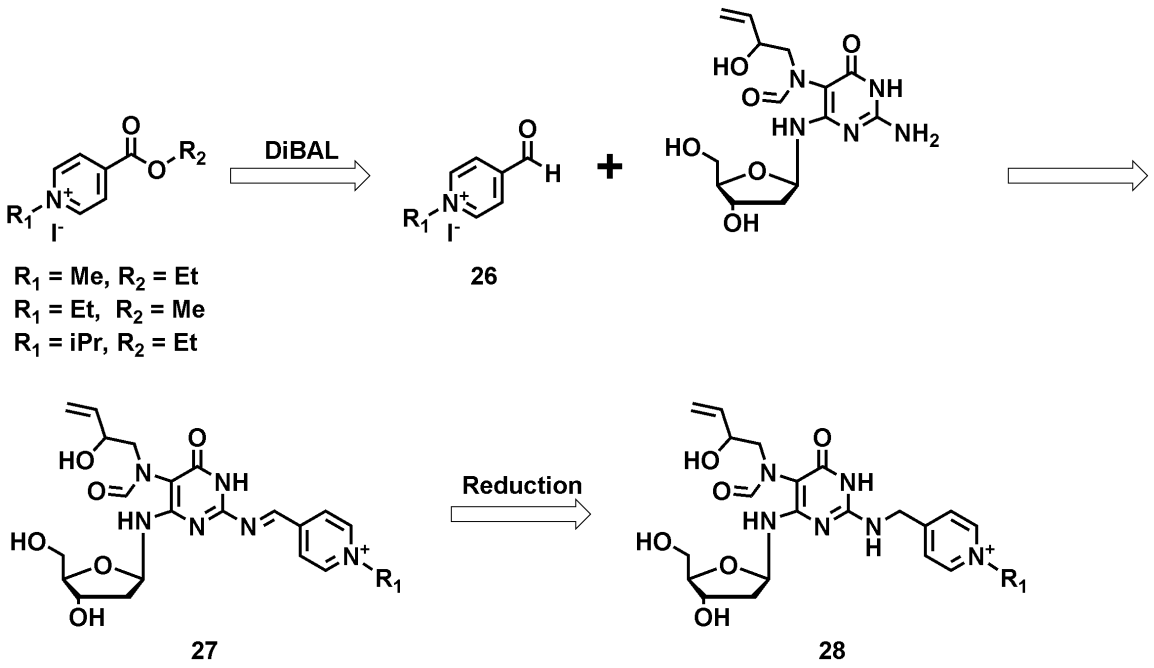
FAPy-dG. Unfortunately, there are no commercially available to reagents attempt these reactions. However, several esters including 4-ethoxycarbonyl-1-methyl-pyridinium iodide, 1-ethyl-4-(methoxycarbonyl)pyridinium iodide, and 4-ethoxycarbonyl-1-isopropyl-pyridinium iodide, which can be reduced to aldehydes, are available from Sigma Aldrich (St. Louis, MO) (**Scheme 7-2**). Specifically, these esters could be reduced with DIBAL at low temperatures (-70 ºC) to afford aldehydes as described previously (**28** in **Scheme 7-2**).[611-613] The synthesized products need to be fully structurally characterized by ^{1}H-NMR and C-NMR, specifically by observing the presence of the protecting group and the aldehyde peak at 9 – 10 ppm and 160 – 170 ppm respectively.

The synthesized aldehydes can then be reacted with EB-FAPy-dG following the same conditions used for *N,N*-dimethylformamide dimethyl acetal protection described in **Section 5.3.1** (**26** in **Scheme 7-2**). Successful derivatization of EB-FAPy-dG can be confirmed by ESI MS, ^{1}H-NMR, and ^{13}C-NMR. The derivatives (**26** in **Scheme 7-2**) would then be purified by HPLC (**Section 5.2**), and the pure product used to develop a new nanoLC-ESI^{+}-MS/MS assay. Since the derivatized EB-FAPy-dG should be less polar than the unprotected nucleoside, more common column packing material such as XDB C18 or SB-300 C18 material could be investigated. If the derivatized EB-FAPy-dG is not stable enough to purify by column or HPLC, the *N,N*-dimethylformimidamide protecting group can be reduced to *N,N*-dimethylmethanediamine by sodium borohydride or sodium cyanoborohydride (**27** in **Scheme 7-2**).

To confirm that the novel aldehyde(s) (**26** in **Scheme 7-2**) can derivatize EB-FAPy-dG formed *in vitro* or *in vivo*, EB-FAPy-dG (**Section 5.3.2**) can be reacted with 20x excess reagent and analyzed by the developed assay. Successful detection of the derivatized EB-

FAPy-dG would confirm that the DNA digestion and enrichment procedure successfully removes any interfering contaminates. SPE recovery after derivatization, synthesis and purification of an isotopically-labeled internal standard, and validation of the assay could then be performed to complete assay development before attempting to quantify EB-FAPy-dG formation in animal or humans exposed to BD.

Scheme 7-2: Synthetic strategy to derivatize the exocyclic amine of EB-FAPy-dG to improve adduct sensitivity.



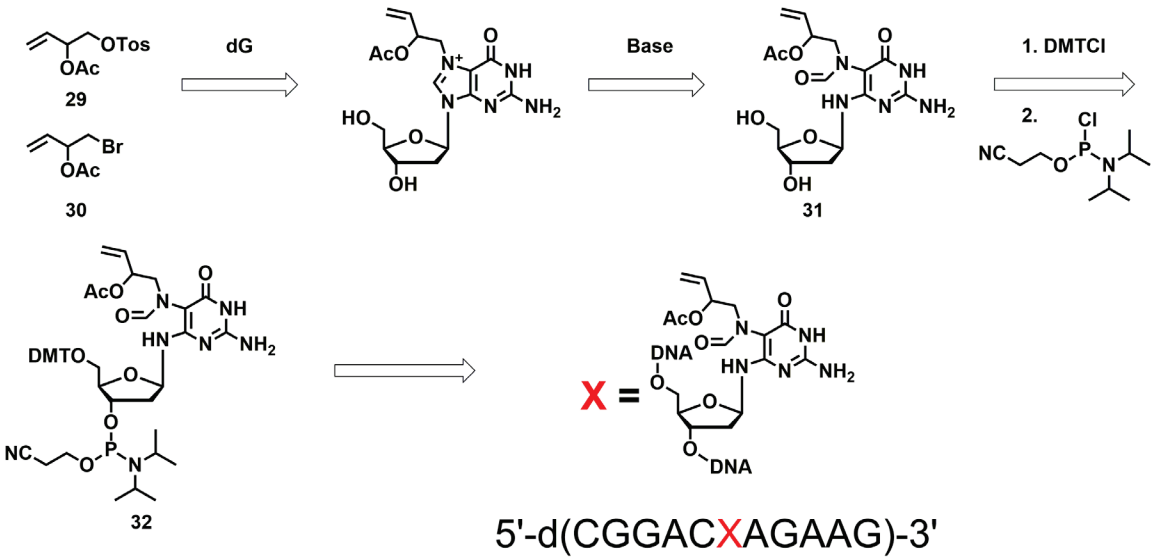
7.7. Site-specific incorporation of EB-FAPy-dG in DNA for structural and thermodynamics studies

Previous studies with site-specific incorporated AFB_1-FAPy-dG, NM-FAPy-dG, and Me-FAPy-dG lesions in duplex DNA have revealed characteristic physical properties. For example, all three lesions in duplex DNA showed a biphasic thermal melting profile,[231, 232, 245] with the β-anomer of the FAPy-dG lesion significantly stabilizing duplex DNA.[232] Furthermore, synthetic oligos containing a NM-FAPy-dG lesion exist in an anomeric ratio of 1:1 (α : β), but will slowly equilibrate to approximately 80% β-anomer.[232] Finally, the equilibrium of the *E* and *Z* formamide isomers of FAPy-dG is controlled by the neighboring nucleotide and subsequent major groove hydrogen bonding interactions.[614] However, we are unable to synthesize DNA oligomers containing an EB-FAPy-dG lesion using EB-FAPy-dG nucleoside prepared in our work (**Scheme 5-1**) due to the presence of multiple secondary alcohols. Therefore, to investigate the biophysical properties of EB-FAPy-dG similar to every other FAPy-dG lesion, we will have to modify our synthetic strategy.

To synthesize the EB-FAPy-dG phosphoramidite, we need to modify our scheme to alkylate protected dG (**3**, **Figure 5.2**) with an alcohol protected EB analog. Two potential EB analogs we could synthesize include 1-(tosyloxy)but-3-en-2-yl acetate and 1-bromobut-3-en-2-yl acetate synthesized from 4-vinyl-1,3-dioxolan-2-one or 1-bromobut-3-en-2-ol starting material respectively (**29** and **30** in **Scheme 7-3** respectively). Successful alkylation and base catalyzed imidazole-ring opening to **31** can be confirmed by mass spec infusion, ^{1}H-NMR, and ^{13}C-NMR. After obtaining the alcohol-protected EB-FAPy-dG nucleoside, the phosphoramidite can be synthesized following previously developed two-step synthetic schemes (**32** in **Scheme 7-3**).[615] Finally, the EB-FAPy-dG adduct can be site-

specifically incorporated into DNA using solid phase synthesis. We will prepare synthetic oligomer (5′-d(CGGAC**X**AGAAG)-3′), where **X** = EB-FAPy-dG phosphoramidite, which was developed for previous NMR chemical perturbation analysis (**Scheme 7-3**).[616]

Scheme 7-3: Synthetic strategy to site-specifically incorporate an EB-FAPy-dG adduct into an 11-mer DNA oligo for structural studies.



After successfully incorporating the EB-FAPy-dG lesion into a synthetic oligonucleotide, we can perform DNA thermal stability studies to calculate the change in melting temperature (ΔT_m) caused by the lesion. Specifically, a solution of unsubstituted DNA oligos 5′-d(CGGACGAGAAG)-3′ and 5′-d(CTTCTTGTCCG)-3’ will be heated from 5 to 90 ºC at a rate of 1.0 ºC/min while monitoring the UV absorbance at 260 nm. The T_m value of the 12-mer duplex can be determined by taking the first derivative of the UV absorbance vs temperature curve. When the EB-FAPy-dG-containing oligo 5′-d(CGGAC**X**AGAAG)-3′ is investigated, we expected to observe two T_m values corresponding to the α and β anomer similar to AFB_1-FAPy-dG and Me-FAPy-dG lesions.[231, 245] Furthermore, we also expect one of the T_m values will be only slightly stabilizing compared to the control (α-anomer, ΔT_m = 1 – 5 ºC), while the other T_m value will be significantly stabilizing compared to the control (β-anomer, ΔT_m = 10 – 15 ºC) as observed with the Me-FAPy-dG lesion.

The synthesized 5′-d(CGGAC**X**AGAAG)-3′ oligo can also be analyzed by high field NMR (600 – 800 mHz) ^{1}H-NMR, NOESY and DQF-COSY experiments to provide structural information to support the thermal melting temperatures. Both NOESY and COSY data can be used to assign the base, amino, and imino protons, allowing chemical shift perturbations to be detected. These chemical shifts will allow us to determine the positioning and hydrogen bond interactions of the 2-hydroxybut-3-en-1-yl constituent, the ratio of furanose to pyranose anomerization, and how the EB-FAPy adduct effects base stacking. Furthermore, sequence effects can be determined by changing the nucleotide immediately downstream the EB-FAPy-dG adduct and analyzing the oligo by HMQC. Specifically, the ratio of the formyl proton can be used to calculate the *E* and *Z* isomer ratio

of the formamide group, providing additional information about the orientation of the adduct.

The availability of the 12-mer oligo 5′-d(CGGAC**X**AGAAG)-3′ would also allow us to confirm the anomerization rate of EB-FAPy-dG *in-vitro*. As described previously, the lesion-containing strand can be labeled with ^{32}P and annealed with the complementary strand, followed by adding the FPG enzyme (180 nmol) and obtaining aliquots 1, 12, 24, 48, 72, and 96 hours post annealing.[232] FPG incision of the EB-FAPy-dG α-anomer can be heated at 90 ºC to hydrolyze the DNA, followed by analysis with SDS-PAGE. We would expect to observe nearly 50% of the duplex incised after a short period of heating, with the nearly all the EB-FAPy-dG lesion anomerized to the β-anomer. Failure to observe loss of the α-anomer would indicate that EB-FAPy-dG is more stable than NM-FAPy-dG, and further NMR analysis would be required to understand the hydrogen bonds that stabilize the lesion.

7.8. Polymerase bypass of EB-FAPy-dG adduct

As described in **Section 1.3.2**, Me-FAPy-dG lesions block high fidelity eukaryotic DNA polymerases at the insertion or the extension step, and TLS polymerases bypass the lesion in an error-prone manner.[234] Furthermore, AFB_1-FAPy-dG lesions lead to G → T transversion mutations in human hepatocytes.[249, 250] Successful synthesis of a lesion-containing 18-mer oligonuclotide (5′-TCAT**X**GAATCCTTCCCCC-3', where **X** = EB-FAPy-dG) would allow us to investigate "standing start" and "running start" primer extension studies and measure steady state kinetics of nucleotide insertion across the lesion as described previously.[559]

Briefly, for "standing start" experiments an annealed mixture of ^{32}P-labeled 13mer (5′-GGGGGAAGGATTC-3′) and lesion-containing 18-mer template will be incubated with a human DNA polymerase (hPol β, hPol η, hPol ι, hPol ζ and hPol κ). ADD A SCHEME The primer extension reaction will be initiated by adding a dNTP mix, with aliquots obtained and quenched every 10 min for 1 h. The primer extension products at every time point can be resolved by gel electrophoresis and observed by phosphorimaging. For running start reactions, the 13-mer primer can be replaced by a shorter 9-mer primer (5'-GGGGGAAGG-3') and the experiment above repeated. Additionally, cooperativity of TLS polymerases in bypassing EB-FAPy will be investigated by incubating multiple enzymes with the annealed DNA and repeating the experiments described above. Based on previous reports for Me-FAPy-dG,[234] we anticipate hPol β to be fully blocked by the lesion, while the TLS polymerases will be able to bypass the lesion.

In order to investigate the fidelity of nucleotide insertion across the EB-FAPy-dG adduct, the "standing start" experiment can be repeated by increasing the concentration of

an individual dNTPs over a specific time period. Steady-state kinetic parameters (k_{cat} and K_m), catalytic specificity constants (k_{cat}/K_m), and the misinsertion frequency of nucleotide insertion can be calculated from the experiments above as described in detail previously.[617] We expect to observe several TLS polymerases inefficiently insert all four nucleotides across the lesion, showing that EB-FAPy-dG is mutagenic.

7.9. Elucidating the repair enzymes involved in excising EB-FAPy-dG

As observed in **Section 5.3.4**, we observed a slight difference in EB-FAPy-dG formation in MEF NEIL1$^{-/-}$ cells as compared to the wild type NEIL1$^{+/+}$ cells, indicating that the BER enzyme NEIL1 may contribute to the repair of EB-FAPy-dG. To fully elucidate the enzymes that recognize EB-FAPy-dG as a substrate, we will incubate the ^{32}P-labeled 12mer lesion-containing oligo with a commercially available BER enzyme (OGG1, NTHL1, NEIL1, NEIL2, and NEIL3). Aliquots will be obtained 1, 2, 4, 8, and 24 hours and analyzed as described in **Section 6.8**. Observation of a smaller 5-mer DNA fragments will provide evidence of repair. Based on already confirmed substrate selectivity (**Table 1.2**), we expect one of the NEIL enzymes will efficiently excise EB-FAPy-dG. If no BER enzyme is able to excise EB-FAPy-dG *in-vitro*, we can investigate NER repair of the adduct following a different assay.[618]

Alternatively, an EB-FAPy-dG lesion can be site-specifically incorporated into a biotin-labeled DNA oligo and incubated in a cellular lysate. Subsequent enrichment using biotin-labeled beads would isolate the oligo and any proteins bound to the strand. The enriched proteins can then be removed from the DNA under denaturing conditions, and the proteins identified by mass spectrometry-based proteomics technique described in **Section 2.2**. As a control, unmodified DNA can be incubated in the same lysate and the proteins analyzed as described above. Identification of any DNA repair enzymes could provide insight to the mechanism of EB-FAPy-dG repair, and lead to a directed *in-vitro* investigation described above.

BIBLIOGRAPHY

1. Alberts, B. (2003) DNA replication and recombination, *Nature 421*, 431-435.
2. Parker, M. W., Botchan, M. R., and Berger, J. M. (2017) Mechanisms and regulation of DNA replication initiation in eukaryotes, *Crit Rev. Biochem. Mol. Biol 52*, 107-144.
3. Gates, K. S. (2009) An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals, *Chem. Res. Toxicol 22*, 1747-1760.
4. Lindahl, T. (1993) Instability and decay of the primary structure of DNA, *Nature 362*, 709-715.
5. Drablos, F., Feyzi, E., Aas, P. A., Vaagbo, C. B., Kavli, B., Bratlie, M. S., Pena-Diaz, J., Otterlei, M., Slupphaug, G., and Krokan, H. E. (2004) Alkylation damage in DNA and RNA--repair mechanisms and medical significance, *DNA Repair (Amst) 3*, 1389-1407.
6. Fu, D., Calvo, J. A., and Samson, L. D. (2012) Balancing repair and tolerance of DNA damage caused by alkylating agents, *Nat. Rev. Cancer 12*, 104-120.
7. Hoeijmakers, J. H. (2001) Genome maintenance mechanisms for preventing cancer, *Nature 411*, 366-374.
8. Ghosal, G., and Chen, J. (2013) DNA damage tolerance: a double-edged sword guarding the genome, *Transl. Cancer Res 2*, 107-129.
9. Sinha, R. P., and Hader, D. P. (2002) UV-induced DNA damage and repair: a review, *Photochem. Photobiol. Sci 1*, 225-236.
10. Rastogi, R. P., Richa, Kumar, A., Tyagi, M. B., and Sinha, R. P. (2010) Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair, *J. Nucleic Acids 2010*, 592980.
11. Cooke, M. S., Evans, M. D., Dizdaroglu, M., and Lunec, J. (2003) Oxidative DNA damage: mechanisms, mutation, and disease, *FASEB J 17*, 1195-1214.
12. Farmer, P. B., Singh, R., Kaur, B., Sram, R. J., Binkova, B., Kalina, I., Popov, T. A., Garte, S., Taioli, E., Gabelova, A., and Cebulska-Wasilewska, A. (2003) Molecular epidemiology studies of carcinogenic environmental pollutants. Effects of polycyclic aromatic hydrocarbons (PAHs) in environmental pollution on exogenous and oxidative DNA damage, *Mutat Res 544*, 397-402.
13. Xue, W., and Warshawsky, D. (2005) Metabolic activation of polycyclic and heterocyclic aromatic hydrocarbons and DNA damage: a review, *Toxicol Appl Pharmacol 206*, 73-93.
14. De Bont, R., and van Larebeke, N. (2004) Endogenous DNA damage in humans: a review of quantitative data, *Mutagenesis 19*, 169-185.
15. Shapiro, R., and Klein, R. S. (1966) The deamination of cytidine and cytosine by acidic buffer solutions. Mutagenic implications, *Biochemistry 5*, 2358-2362.
16. Shapiro, R., and Danzig, M. (1972) Acidic hydrolysis of deoxycytidine and deoxyuridine derivatives. The general mechanism of deoxyribonucleoside hydrolysis, *Biochemistry 11*, 23-29.

17. Lindahl, T., and Nyberg, B. (1974) Heat-induced deamination of cytosine residues in deoxyribonucleic acid, *Biochemistry 13*, 3405-3410.
18. Frederico, L. A., Kunkel, T. A., and Shaw, B. R. (1990) A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy, *Biochemistry 29*, 2532-2537.
19. Garrett, E. R., and Tsau, J. (1972) Solvolyses of cytosine and cytidine, *J. Pharm. Sci 61*, 1052-1061.
20. Karran, P., and Lindahl, T. (1980) Hypoxanthine in deoxyribonucleic acid: generation by heat-induced hydrolysis of adenine residues and release in free form by a deoxyribonucleic acid glycosylase from calf thymus, *Biochemistry 19*, 6005-6011.
21. Lindahl, T., and Nyberg, B. (1972) Rate of depurination of native deoxyribonucleic acid, *Biochemistry 11*, 3610-3618.
22. Zoltewicz, J. A., Clark, D. F., Sharpless, T. W., and Grahe, G. (1970) Kinetics and mechanism of the acid-catalyzed hydrolysis of some purine nucleosides, *J. Am. Chem. Soc 92*, 1741-1749.
23. Lindahl, T., and Andersson, A. (1972) Rate of chain breakage at apurinic sites in double-stranded deoxyribonucleic acid, *Biochemistry 11*, 3618-3623.
24. Schroeder, G. K., Lad, C., Wyman, P., Williams, N. H., and Wolfenden, R. (2006) The time required for water attack at the phosphorus atom of simple phosphodiesters and of DNA, *Proc. Natl. Acad. Sci. U. S. A 103*, 4052-4055.
25. Cassano, A. G., Anderson, V. E., and Harris, M. E. (2002) Evidence for direct attack by hydroxide in phosphodiester hydrolysis, *J. Am. Chem. Soc 124*, 10964-10965.
26. Iche-Tarrat, N., Barthelat, J. C., Rinaldi, D., and Vigroux, A. (2005) Theoretical studies of the hydroxide-catalyzed P-O cleavage reactions of neutral phosphate triesters and diesters in aqueous solution: examination of the changes induced by H/Me substitution, *J. Phys. Chem. B 109*, 22570-22580.
27. Crine, P., and Verly, W. G. (1976) A study of DNA spontaneous degradation, *Biochim. Biophys. Acta 442*, 50-57.
28. Dutta, S., Chowdhury, G., and Gates, K. S. (2007) Interstrand cross-links generated by abasic sites in duplex DNA, *J. Am. Chem. Soc 129*, 1852-1853.
29. Quinones, J. L., Thapar, U., Yu, K., Fang, Q., Sobol, R. W., and Demple, B. (2015) Enzyme mechanism-based, oxidative DNA-protein cross-links formed with DNA polymerase beta in vivo, *Proc. Natl. Acad. Sci. U. S. A 112*, 8602-8607.
30. Quinones, J. L., and Demple, B. (2016) When DNA repair goes wrong: BER-generated DNA-protein crosslinks to oxidative lesions, *DNA Repair (Amst) 44*, 103-109.
31. Cadet, J., and Wagner, J. R. (2013) DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation, *Cold Spring Harb. Perspect. Biol 5*.
32. Burney, S., Caulfield, J. L., Niles, J. C., Wishnok, J. S., and Tannenbaum, S. R. (1999) The chemistry of DNA damage from nitric oxide and peroxynitrite, *Mutat. Res 424*, 37-49.
33. Szabo, C., and Ohshima, H. (1997) DNA damage induced by peroxynitrite: subsequent biological effects, *Nitric. Oxide 1*, 373-385.
34. Cadet, J., Douki, T., and Ravanat, J. L. (2010) Oxidatively generated base damage to cellular DNA, *Free Radic. Biol. Med 49*, 9-21.

35. Halliwell, B., and Aruoma, O. I. (1991) DNA damage by oxygen-derived species. Its mechanism and measurement in mammalian systems, *FEBS Lett 281*, 9-19.
36. Breen, A. P., and Murphy, J. A. (1995) Reactions of oxyl radicals with DNA, *Free Radic. Biol. Med 18*, 1033-1077.
37. Greenberg, M. M. (1998) Investigating nucleic acid damage processes via independent generation of reactive intermediates, *Chem. Res. Toxicol 11*, 1235-1248.
38. Pogozelski, W. K., and Tullius, T. D. (1998) Oxidative Strand Scission of Nucleic Acids: Routes Initiated by Hydrogen Abstraction from the Sugar Moiety, *Chem. Rev 98*, 1089-1108.
39. Balasubramanian, B., Pogozelski, W. K., and Tullius, T. D. (1998) DNA strand breaking by the hydroxyl radical is governed by the accessible surface areas of the hydrogen atoms of the DNA backbone, *Proc. Natl. Acad. Sci. U. S. A 95*, 9738-9743.
40. Burrows, C. J., and Muller, J. G. (1998) Oxidative Nucleobase Modifications Leading to Strand Scission, *Chem. Rev 98*, 1109-1152.
41. Gajewski, E., Rao, G., Nackerdien, Z., and Dizdaroglu, M. (1990) Modification of DNA bases in mammalian chromatin by radiation-generated free radicals, *Biochemistry 29*, 7876-7882.
42. Wallace, S. S. (2002) Biological consequences of free radical-damaged DNA bases, *Free Radic. Biol. Med 33*, 1-14.
43. Jovanovic, S. V., and Simic, M. G. (1986) Mechanism of OH radical reactions with thymine and uracil derivatives, *J. Am. Chem. Soc 108*, 5968-5972.
44. Douki, T., Ravanat, J. L., Pouget, J. P., Testard, I., and Cadet, J. (2006) Minor contribution of direct ionization to DNA base damage inducedby heavy ions, *Int. J. Radiat. Biol 82*, 119-127.
45. Pouget, J. P., Frelon, S., Ravanat, J. L., Testard, I., Odin, F., and Cadet, J. (2002) Formation of modified DNA bases in cells exposed either to gamma radiation or to high-LET particles, *Radiat. Res 157*, 589-595.
46. Candeias, L. P., and Steenken, S. (2000) Reaction of HO* with guanine derivatives in aqueous solution: formation of two different redox-active OH-adduct radicals and their unimolecular transformation reactions. Properties of G(-H)*, *Chemistry 6*, 475-484.
47. Kasai, H., and Nishimura, S. (1983) Hydroxylation of the C-8 position of deoxyguanosine by reducing agents in the presence of oxygen, *Nucleic Acids Symp. Ser*, 165-167.
48. van den, A. E., Lutgerink, J. T., Lafleur, M. V., Joenje, H., and Retel, J. (1994) The formation of one-G deletions as a consequence of single-oxygen-induced DNA damage, *Mutat. Res 309*, 45-52.
49. Bellon, S., Ravanat, J. L., Gasparutto, D., and Cadet, J. (2002) Cross-linked thymine-purine base tandem lesions: synthesis, characterization, and measurement in gamma-irradiated isolated DNA, *Chem. Res. Toxicol 15*, 598-606.
50. Box, H. C., Budzinski, E. E., Dawidzik, J. B., Wallace, J. C., and Iijima, H. (1998) Tandem lesions and other products in X-irradiated DNA oligomers, *Radiat. Res 149*, 433-439.

51. Luo, W., Muller, J. G., Rachlin, E. M., and Burrows, C. J. (2000) Characterization of spiroiminodihydantoin as a product of one-electron oxidation of 8-Oxo-7,8-dihydroguanosine, *Org. Lett 2*, 613-616.
52. Luo, W., Muller, J. G., Rachlin, E. M., and Burrows, C. J. (2001) Characterization of hydantoin products from one-electron oxidation of 8-oxo-7,8-dihydroguanosine in a nucleoside model, *Chem. Res. Toxicol 14*, 927-938.
53. Ye, Y., Muller, J. G., Luo, W., Mayne, C. L., Shallop, A. J., Jones, R. A., and Burrows, C. J. (2003) Formation of 13C-, 15N-, and 18O-labeled guanidinohydantoin from guanosine oxidation with singlet oxygen. Implications for structure and mechanism, *J. Am. Chem. Soc 125*, 13926-13927.
54. Matter, B., Malejka-Giganti, D., Csallany, A. S., and Tretyakova, N. (2006) Quantitative analysis of the oxidative DNA lesion, 2,2-diamino-4-(2-deoxy-beta-D-erythro-pentofuranosyl)amino]-5(2H)-oxazolone (oxazolone), in vitro and in vivo by isotope dilution-capillary HPLC-ESI-MS/MS, *Nucleic Acids Res 34*, 5449-5460.
55. Neeley, W. L., and Essigmann, J. M. (2006) Mechanisms of formation, genotoxicity, and mutation of guanine oxidation products, *Chem. Res. Toxicol 19*, 491-505.
56. Cadet, J., Ravanat, J. L., TavernaPorro, M., Menoni, H., and Angelov, D. (2012) Oxidatively generated complex DNA damage: tandem and clustered lesions, *Cancer Lett 327*, 5-15.
57. Tallman, K. A., and Greenberg, M. M. (2001) Oxygen-dependent DNA damage amplification involving 5,6-dihydrothymidin-5-yl in a structurally minimal system, *J. Am. Chem. Soc 123*, 5181-5187.
58. Budzinski, E. E., Dawidzik, J. D., Wallace, J. C., Freund, H. G., and Box, H. C. (1995) The radiation chemistry of d(CpGpTpA) in the presence of oxygen, *Radiat. Res 142*, 107-109.
59. Brooks, P. J., Wise, D. S., Berry, D. A., Kosmoski, J. V., Smerdon, M. J., Somers, R. L., Mackie, H., Spoonde, A. Y., Ackerman, E. J., Coleman, K., Tarone, R. E., and Robbins, J. H. (2000) The oxidative DNA lesion 8,5'-(S)-cyclo-2'-deoxyadenosine is repaired by the nucleotide excision repair pathway and blocks gene expression in mammalian cells, *J. Biol. Chem 275*, 22355-22362.
60. Dizdaroglu, M., Jaruga, P., and Rodriguez, H. (2001) Identification and quantification of 8,5'-cyclo-2'-deoxy-adenosine in DNA by liquid chromatography/ mass spectrometry, *Free Radic. Biol. Med 30*, 774-784.
61. Yoon, J. H., Lee, C. S., O'Connor, T. R., Yasui, A., and Pfeifer, G. P. (2000) The DNA damage spectrum produced by simulated sunlight, *J. Mol. Biol 299*, 681-693.
62. You, Y. H., Szabo, P. E., and Pfeifer, G. P. (2000) Cyclobutane pyrimidine dimers form preferentially at the major p53 mutational hotspot in UVB-induced mouse skin tumors, *Carcinogenesis 21*, 2113-2117.
63. Clingen, P. H., Arlett, C. F., Roza, L., Mori, T., Nikaido, O., and Green, M. H. (1995) Induction of cyclobutane pyrimidine dimers, pyrimidine(6-4)pyrimidone photoproducts, and Dewar valence isomers by natural sunlight in normal human mononuclear cells, *Cancer Res 55*, 2245-2248.

64. Mitchell, D. L., and Rosenstein, B. S. (1987) The use of specific radioimmunoassay to determine action spectra for the photolysis of (6-4) photoproducts, *Photochem. Photobiol 45*, 781-786.
65. Lomax, M. E., Folkes, L. K., and O'Neill, P. (2013) Biological consequences of radiation-induced DNA damage: relevance to radiotherapy, *Clin. Oncol. (R. Coll. Radiol.) 25*, 578-585.
66. Shrivastav, N., Li, D., and Essigmann, J. M. (2010) Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation, *Carcinogenesis 31*, 59-70.
67. Beranek, D. T. (1990) Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents, *Mutat. Res 231*, 11-30.
68. Loechler, E. L. (1994) A violation of the Swain-Scott principle, and not SN1 versus SN2 reaction mechanisms, explains why carcinogenic alkylating agents can form different proportions of adducts at oxygen versus nitrogen in DNA, *Chem. Res. Toxicol 7*, 277-280.
69. Lu, X., Heilman, J. M., Blans, P., and Fishbein, J. C. (2005) The structure of DNA dictates purine atom site selectivity in alkylation by primary diazonium ions, *Chem. Res. Toxicol 18*, 1462-1470.
70. Pearson, R. (1963) Hard and soft acids and bases, *J. Am. Chem. Soc 85*, 6.
71. Perrino, F. W., Blans, P., Harvey, S., Gelhaus, S. L., McGrath, C., Akman, S. A., Jenkins, G. S., LaCourse, W. R., and Fishbein, J. C. (2003) The N2-ethylguanine and the O6-ethyl- and O6-methylguanine lesions in DNA: contrasting responses from the "bypass" DNA polymerase eta and the replicative DNA polymerase alpha, *Chem. Res. Toxicol 16*, 1616-1623.
72. Wilds, C. J., Noronha, A. M., Robidoux, S., and Miller, P. S. (2004) Mispair-aligned N3T-alkyl-N3T interstrand cross-linked DNA: synthesis and characterization of duplexes with interstrand cross-links of variable lengths, *J. Am. Chem. Soc 126*, 9257-9265.
73. Spratt, T. E., and Campbell, C. R. (1994) Synthesis of oligodeoxynucleotides containing analogs of O6-methylguanine and reaction with O6-alkylguanine-DNA alkyltransferase, *Biochemistry 33*, 11364-11371.
74. Singer, B., Kroger, M., and Carrano, M. (1978) O2- and O4-alkyl pyrimidine nucleosides: stability of the glycosyl bond and of the alkyl group as a function of pH, *Biochemistry 17*, 1246-1250.
75. Sowers, L. C., Sedwick, W. D., and Shaw, B. R. (1989) Hydrolysis of N3-methyl-2'-deoxycytidine: model compound for reactivity of protonated cytosine residues in DNA, *Mutat. Res 215*, 131-138.
76. Lawley, P. D., and Warren, W. (1976) Removal of minor methylation products 7-methyladenine and 3-methylguanine from DNA of Escherichia coli treated with dimethyl sulphate, *Chem. Biol. Interact 12*, 211-220.
77. Lawley, P. D., and BROOKES, P. (1963) FURTHER STUDIES ON THE ALKYLATION OF NUCLEIC ACIDS AND THEIR CONSTITUENT NUCLEOTIDES, *Biochem. J 89*, 127-138.
78. Gates, K. S., Nooner, T., and Dutta, S. (2004) Biologically relevant chemical reactions of N7-alkylguanine residues in DNA, *Chem. Res. Toxicol 17*, 839-856.

79. Chetsanga, C. J., and Makaroff, C. (1982) Alkaline opening of imidazole ring of 7-methylguanosine. 2. Further studies on reaction mechanisms and products, *Chem. Biol. Interact 41*, 235-249.
80. Chetsanga, C. J., Bearie, B., and Makaroff, C. (1982) Alkaline opening of imidazole ring of 7-methylguanosine. 1. Analysis of the resulting pyrimidine derivatives, *Chem. Biol. Interact 41*, 217-233.
81. Barlow, T., Takeshita, J., and Dipple, A. (1998) Deamination and Dimroth rearrangement of deoxyadenosine-styrene oxide adducts in DNA, *Chem. Res. Toxicol 11*, 838-845.
82. Macon, J. B., and Wolfenden, R. (1968) 1-Methyladenosine. Dimroth rearrangement and reversible reduction, *Biochemistry 7*, 3453-3458.
83. Yi, C., and He, C. (2013) DNA repair by reversal of DNA damage, *Cold Spring Harb. Perspect. Biol 5*, a012575.
84. Muller, M., and Carell, T. (2009) Structural biology of DNA photolyases and cryptochromes, *Curr. Opin. Struct. Biol 19*, 277-285.
85. Brettel, K., and Byrdin, M. (2010) Reaction mechanisms of DNA photolyase, *Curr. Opin. Struct. Biol 20*, 693-701.
86. Pegg, A. E. (2011) Multifaceted roles of alkyltransferase and related proteins in DNA repair, DNA damage, resistance to chemotherapy, and research tools, *Chem. Res. Toxicol 24*, 618-639.
87. Daniels, D. S., Woo, T. T., Luu, K. X., Noll, D. M., Clarke, N. D., Pegg, A. E., and Tainer, J. A. (2004) DNA binding and nucleotide flipping by the human DNA repair protein AGT, *Nat. Struct. Mol. Biol 11*, 714-720.
88. Falnes, P. O., Johansen, R. F., and Seeberg, E. (2002) AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli, *Nature 419*, 178-182.
89. Falnes, P. O., Klungland, A., and Alseth, I. (2007) Repair of methyl lesions in DNA and RNA by oxidative demethylation, *Neuroscience 145*, 1222-1232.
90. Sedgwick, B. (2004) Repairing DNA-methylation damage, *Nat. Rev. Mol. Cell Biol 5*, 148-157.
91. Yi, C., Jia, G., Hou, G., Dai, Q., Zhang, W., Zheng, G., Jian, X., Yang, C. G., Cui, Q., and He, C. (2010) Iron-catalysed oxidation intermediates captured in a DNA repair dioxygenase, *Nature 468*, 330-333.
92. Yi, C., Yang, C. G., and He, C. (2009) A non-heme iron-mediated chemical demethylation in DNA and RNA, *Acc. Chem. Res 42*, 519-529.
93. Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., and Sedgwick, B. (2002) Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage, *Nature 419*, 174-178.
94. Delaney, J. C., and Essigmann, J. M. (2004) Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli, *Proc. Natl. Acad. Sci. U. S. A 101*, 14051-14056.
95. Koivisto, P., Robins, P., Lindahl, T., and Sedgwick, B. (2004) Demethylation of 3-methylthymine in DNA by bacterial and human DNA dioxygenases, *J. Biol. Chem 279*, 40470-40474.
96. Delaney, J. C., Smeester, L., Wong, C., Frick, L. E., Taghizadeh, K., Wishnok, J. S., Drennan, C. L., Samson, L. D., and Essigmann, J. M. (2005) AlkB reverses

etheno DNA lesions caused by lipid oxidation in vitro and in vivo, *Nat. Struct. Mol. Biol 12*, 855-860.
97. Frick, L. E., Delaney, J. C., Wong, C., Drennan, C. L., and Essigmann, J. M. (2007) Alleviation of 1,N6-ethanoadenine genotoxicity by the Escherichia coli adaptive response protein AlkB, *Proc. Natl. Acad. Sci. U. S. A 104*, 755-760.
98. Mishina, Y., Yang, C. G., and He, C. (2005) Direct repair of the exocyclic DNA adduct 1,N6-ethenoadenine by the DNA repair AlkB proteins, *J. Am. Chem. Soc 127*, 14594-14595.
99. Kim, Y. J., and Wilson, D. M., III. (2012) Overview of base excision repair biochemistry, *Curr. Mol. Pharmacol 5*, 3-13.
100. Krokan, H. E., and Bjoras, M. (2013) Base excision repair, *Cold Spring Harb. Perspect. Biol 5*, a012583.
101. Robertson, A. B., Klungland, A., Rognes, T., and Leiros, I. (2009) DNA repair in mammalian cells: Base excision repair: the long and short of it, *Cell Mol. Life Sci 66*, 981-993.
102. Wibley, J. E., Waters, T. R., Haushalter, K., Verdine, G. L., and Pearl, L. H. (2003) Structure and specificity of the vertebrate anti-mutator uracil-DNA glycosylase SMUG1, *Mol. Cell 11*, 1647-1659.
103. Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E., and Tainer, J. A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA, *Nature 384*, 87-92.
104. Huffman, J. L., Sundheim, O., and Tainer, J. A. (2005) DNA base damage recognition and removal: new twists and grooves, *Mutat. Res 577*, 55-76.
105. Mitra, S. (2001) DNA glycosylases: specificity and mechanisms, *Prog. Nucleic Acid Res. Mol. Biol 68*, 189-192.
106. O'Connor, T. R., and Laval, J. (1989) Physical association of the 2,6-diamino-4-hydroxy-5N-formamidopyrimidine-DNA glycosylase of Escherichia coli and an activity nicking DNA at apurinic/apyrimidinic sites, *Proc. Natl. Acad. Sci. U. S. A 86*, 5222-5226.
107. Beard, W. A., Prasad, R., and Wilson, S. H. (2006) Activities and mechanism of DNA polymerase beta, *Methods Enzymol 408*, 91-107.
108. Kubota, Y., Nash, R. A., Klungland, A., Schar, P., Barnes, D. E., and Lindahl, T. (1996) Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein, *EMBO J 15*, 6662-6670.
109. Matsumoto, Y., and Kim, K. (1995) Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair, *Science 269*, 699-702.
110. Singhal, R. K., and Wilson, S. H. (1993) Short gap-filling synthesis by DNA polymerase beta is processive, *J. Biol. Chem 268*, 15906-15911.
111. Sobol, R. W., Horton, J. K., Kuhn, R., Gu, H., Singhal, R. K., Prasad, R., Rajewsky, K., and Wilson, S. H. (1996) Requirement of mammalian DNA polymerase-beta in base-excision repair, *Nature 379*, 183-186.
112. Wei, Y. F., Robins, P., Carter, K., Caldecott, K., Pappin, D. J., Yu, G. L., Wang, R. P., Shell, B. K., Nash, R. A., Schar, P., and . (1995) Molecular cloning and expression of human cDNAs encoding a novel DNA ligase IV and DNA ligase

III, an enzyme active in DNA repair and recombination, *Mol. Cell Biol 15*, 3206-3216.
113. Frosina, G., Fortini, P., Rossi, O., Carrozzino, F., Raspaglio, G., Cox, L. S., Lane, D. P., Abbondandolo, A., and Dogliotti, E. (1996) Two pathways for base excision repair in mammalian cells, *J. Biol. Chem 271*, 9573-9578.
114. Kim, K., Biade, S., and Matsumoto, Y. (1998) Involvement of flap endonuclease 1 in base excision DNA repair, *J. Biol. Chem 273*, 8842-8848.
115. Klungland, A., and Lindahl, T. (1997) Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1), *EMBO J 16*, 3341-3348.
116. Petermann, E., Ziegler, M., and Oei, S. L. (2003) ATP-dependent selection between single nucleotide and long patch base excision repair, *DNA Repair (Amst) 2*, 1101-1114.
117. Couve-Privat, S., Mace, G., Rosselli, F., and Saparbaev, M. K. (2007) Psoralen-induced DNA adducts are substrates for the base excision repair pathway in human cells, *Nucleic Acids Res 35*, 5672-5682.
118. McNeill, D. R., Paramasivam, M., Baldwin, J., Huang, J., Vyjayanti, V. N., Seidman, M. M., and Wilson, D. M., 3rd. (2013) NEIL1 responds and binds to psoralen-induced DNA interstrand crosslinks, *J Biol Chem 288*, 12426-12436.
119. Boyce, R. P., and HOWARD-FLANDERS, P. (1964) RELEASE OF ULTRAVIOLET LIGHT-INDUCED THYMINE DIMERS FROM DNA IN E. COLI K-12, *Proc. Natl. Acad. Sci. U. S. A 51*, 293-300.
120. PETTIJOHN, D., and HANAWALT, P. (1964) EVIDENCE FOR REPAIR-REPLICATION OF ULTRAVIOLET DAMAGED DNA IN BACTERIA, *J. Mol. Biol 9*, 395-410.
121. RASMUSSEN, R. E., and PAINTER, R. B. (1964) EVIDENCE FOR REPAIR OF ULTRA-VIOLET DAMAGED DEOXYRIBONUCLEIC ACID IN CULTURED MAMMALIAN CELLS, *Nature 203*, 1360-1362.
122. SETLOW, R. B., and CARRIER, W. L. (1964) THE DISAPPEARANCE OF THYMINE DIMERS FROM DNA: AN ERROR-CORRECTING MECHANISM, *Proc. Natl. Acad. Sci. U. S. A 51*, 226-231.
123. Scharer, O. D. (2013) Nucleotide excision repair in eukaryotes, *Cold Spring Harb. Perspect. Biol 5*, a012609.
124. Vermeulen, W., and Fousteri, M. (2013) Mammalian transcription-coupled excision repair, *Cold Spring Harb. Perspect. Biol 5*, a012625.
125. Hanawalt, P. C., and Spivak, G. (2008) Transcription-coupled DNA repair: two decades of progress and surprises, *Nat. Rev. Mol. Cell Biol 9*, 958-970.
126. Sugasawa, K., Okamoto, T., Shimizu, Y., Masutani, C., Iwai, S., and Hanaoka, F. (2001) A multistep damage recognition mechanism for global genomic nucleotide excision repair, *Genes Dev 15*, 507-521.
127. Evans, E., Fellows, J., Coffer, A., and Wood, R. D. (1997) Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein, *EMBO J 16*, 625-638.
128. Yokoi, M., Masutani, C., Maekawa, T., Sugasawa, K., Ohkuma, Y., and Hanaoka, F. (2000) The xeroderma pigmentosum group C protein complex XPC-HR23B plays

an important role in the recruitment of transcription factor IIH to damaged DNA, *J. Biol. Chem 275*, 9870-9875.
129. Volker, M., Mone, M. J., Karmakar, P., van Hoffen, A., Schul, W., Vermeulen, W., Hoeijmakers, J. H., van Driel, R., van Zeeland, A. A., and Mullenders, L. H. (2001) Sequential assembly of the nucleotide excision repair factors in vivo, *Mol. Cell 8*, 213-224.
130. Coin, F., Oksenych, V., and Egly, J. M. (2007) Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair, *Mol. Cell 26*, 245-256.
131. Winkler, G. S., Araujo, S. J., Fiedler, U., Vermeulen, W., Coin, F., Egly, J. M., Hoeijmakers, J. H., Wood, R. D., Timmers, H. T., and Weeda, G. (2000) TFIIH with inactive XPD helicase functions in transcription initiation but is defective in DNA repair, *J. Biol. Chem 275*, 4258-4266.
132. Mathieu, N., Kaczmarek, N., and Naegeli, H. (2010) Strand- and site-specific DNA lesion demarcation by the xeroderma pigmentosum group D helicase, *Proc. Natl. Acad. Sci. U. S. A 107*, 17545-17550.
133. Wakasugi, M., and Sancar, A. (1998) Assembly, subunit composition, and footprint of human DNA repair excision nuclease, *Proc. Natl. Acad. Sci. U. S. A 95*, 6669-6674.
134. de Laat, W. L., Appeldoorn, E., Sugasawa, K., Weterings, E., Jaspers, N. G., and Hoeijmakers, J. H. (1998) DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair, *Genes Dev 12*, 2598-2609.
135. Overmeer, R. M., Moser, J., Volker, M., Kool, H., Tomkinson, A. E., van Zeeland, A. A., Mullenders, L. H., and Fousteri, M. (2011) Replication protein A safeguards genome integrity by controlling NER incision events, *J. Cell Biol 192*, 401-415.
136. Li, L., Elledge, S. J., Peterson, C. A., Bales, E. S., and Legerski, R. J. (1994) Specific association between the human DNA repair proteins XPA and ERCC1, *Proc. Natl. Acad. Sci. U. S. A 91*, 5012-5016.
137. You, J. S., Wang, M., and Lee, S. H. (2003) Biochemical analysis of the damage recognition process in nucleotide excision repair, *J. Biol. Chem 278*, 7476-7485.
138. Fagbemi, A. F., Orelli, B., and Scharer, O. D. (2011) Regulation of endonuclease activity in human nucleotide excision repair, *DNA Repair (Amst) 10*, 722-729.
139. Kemp, M. G., Reardon, J. T., Lindsey-Boltz, L. A., and Sancar, A. (2012) Mechanism of release and fate of excised oligonucleotides during nucleotide excision repair, *J. Biol. Chem 287*, 22889-22899.
140. Ogi, T., Limsirichaikul, S., Overmeer, R. M., Volker, M., Takenaka, K., Cloney, R., Nakazawa, Y., Niimi, A., Miki, Y., Jaspers, N. G., Mullenders, L. H., Yamashita, S., Fousteri, M. I., and Lehmann, A. R. (2010) Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells, *Mol. Cell 37*, 714-727.
141. Ogi, T., and Lehmann, A. R. (2006) The Y-family DNA polymerase kappa (pol kappa) functions in mammalian nucleotide-excision repair, *Nat. Cell Biol 8*, 640-642.
142. Moser, J., Kool, H., Giakzidis, I., Caldecott, K., Mullenders, L. H., and Fousteri, M. I. (2007) Sealing of chromosomal DNA nicks during nucleotide excision repair

requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner, *Mol. Cell 27*, 311-323.
143. Davis, A. J., and Chen, D. J. (2013) DNA double strand break repair via non-homologous end-joining, *Transl. Cancer Res 2*, 130-143.
144. Chang, H. H. Y., Pannunzio, N. R., Adachi, N., and Lieber, M. R. (2017) Non-homologous DNA end joining and alternative pathways to double-strand break repair, *Nat. Rev. Mol. Cell Biol*.
145. Mari, P. O., Florea, B. I., Persengiev, S. P., Verkaik, N. S., Bruggenwirth, H. T., Modesti, M., Giglia-Mari, G., Bezstarosti, K., Demmers, J. A., Luider, T. M., Houtsmuller, A. B., and van Gent, D. C. (2006) Dynamic assembly of end-joining complexes requires interaction between Ku70/80 and XRCC4, *Proc. Natl. Acad. Sci. U. S. A 103*, 18597-18602.
146. Yano, K., Morotomi-Yano, K., Wang, S. Y., Uematsu, N., Lee, K. J., Asaithamby, A., Weterings, E., and Chen, D. J. (2008) Ku recruits XLF to DNA double-strand breaks, *EMBO Rep 9*, 91-96.
147. Cary, R. B., Peterson, S. R., Wang, J., Bear, D. G., Bradbury, E. M., and Chen, D. J. (1997) DNA looping by Ku and the DNA-dependent protein kinase, *Proc. Natl. Acad. Sci. U. S. A 94*, 4267-4272.
148. Weterings, E., and van Gent, D. C. (2004) The mechanism of non-homologous end-joining: a synopsis of synapsis, *DNA Repair (Amst) 3*, 1425-1435.
149. Bernstein, N. K., Williams, R. S., Rakovszky, M. L., Cui, D., Green, R., Karimi-Busheri, F., Mani, R. S., Galicia, S., Koch, C. A., Cass, C. E., Durocher, D., Weinfeld, M., and Glover, J. N. (2005) The molecular architecture of the mammalian DNA repair enzyme, polynucleotide kinase, *Mol. Cell 17*, 657-670.
150. Ahel, I., Rass, U., El Khamisy, S. F., Katyal, S., Clements, P. M., McKinnon, P. J., Caldecott, K. W., and West, S. C. (2006) The neurodegenerative disease protein aprataxin resolves abortive DNA ligation intermediates, *Nature 443*, 713-716.
151. Roberts, S. A., Strande, N., Burkhalter, M. D., Strom, C., Havener, J. M., Hasty, P., and Ramsden, D. A. (2010) Ku is a 5'-dRP/AP lyase that excises nucleotide damage near broken ends, *Nature 464*, 1214-1217.
152. Ma, Y., Pannicke, U., Schwarz, K., and Lieber, M. R. (2002) Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination, *Cell 108*, 781-794.
153. Povirk, L. F., Zhou, T., Zhou, R., Cowan, M. J., and Yannone, S. M. (2007) Processing of 3'-phosphoglycolate-terminated DNA double strand breaks by Artemis nuclease, *J. Biol. Chem 282*, 3547-3558.
154. Perry, J. J., Yannone, S. M., Holden, L. G., Hitomi, C., Asaithamby, A., Han, S., Cooper, P. K., Chen, D. J., and Tainer, J. A. (2006) WRN exonuclease structure and molecular mechanism imply an editing role in DNA end processing, *Nat. Struct. Mol. Biol 13*, 414-422.
155. Li, S., Kanno, S., Watanabe, R., Ogiwara, H., Kohno, T., Watanabe, G., Yasui, A., and Lieber, M. R. (2011) Polynucleotide kinase and aprataxin-like forkhead-associated protein (PALF) acts as both a single-stranded DNA endonuclease and a single-stranded DNA 3' exonuclease and can participate in DNA end joining in a biochemical system, *J. Biol. Chem 286*, 36368-36377.

156. Li, G. M. (2008) Mechanisms and functions of DNA mismatch repair, *Cell Res 18*, 85-98.
157. Jiricny, J. (2006) The multifaceted mismatch-repair system, *Nat. Rev. Mol. Cell Biol 7*, 335-346.
158. Kunkel, T. A., and Erie, D. A. (2015) Eukaryotic Mismatch Repair in Relation to DNA Replication, *Annu. Rev. Genet 49*, 291-313.
159. Schofield, M. J., and Hsieh, P. (2003) DNA mismatch repair: molecular mechanisms and biological function, *Annu. Rev. Microbiol 57*, 579-608.
160. Blackwell, L. J., Wang, S., and Modrich, P. (2001) DNA chain length dependence of formation and dynamics of hMutSalpha.hMutLalpha.heteroduplex complexes, *J. Biol. Chem 276*, 33233-33240.
161. Plotz, G., Raedle, J., Brieger, A., Trojan, J., and Zeuzem, S. (2002) hMutSalpha forms an ATP-dependent complex with hMutLalpha and hMutLbeta on DNA, *Nucleic Acids Res 30*, 711-718.
162. Lin, Y. L., Shivji, M. K., Chen, C., Kolodner, R., Wood, R. D., and Dutta, A. (1998) The evolutionarily conserved zinc finger motif in the largest subunit of human replication protein A is required for DNA replication and mismatch repair but not for nucleotide excision repair, *J. Biol. Chem 273*, 1453-1461.
163. Ramilo, C., Gu, L., Guo, S., Zhang, X., Patrick, S. M., Turchi, J. J., and Li, G. M. (2002) Partial reconstitution of human DNA mismatch repair in vitro: characterization of the role of human replication protein A, *Mol. Cell Biol 22*, 2037-2046.
164. Genschel, J., Bazemore, L. R., and Modrich, P. (2002) Human exonuclease I is required for 5' and 3' mismatch repair, *J. Biol. Chem 277*, 13302-13311.
165. Longley, M. J., Pierce, A. J., and Modrich, P. (1997) DNA polymerase delta is required for human mismatch repair in vitro, *J. Biol. Chem 272*, 10917-10921.
166. Hoeijmakers, J. H. (2009) DNA damage, aging, and cancer, *N. Engl. J. Med 361*, 1475-1485.
167. Valko, M., Izakovic, M., Mazur, M., Rhodes, C. J., and Telser, J. (2004) Role of oxygen radicals in DNA damage and cancer incidence, *Mol. Cell Biochem 266*, 37-56.
168. De Flora, S., Izzotti, A., Randerath, K., Randerath, E., Bartsch, H., Nair, J., Balansky, R., van Schooten, F., Degan, P., Fronza, G., Walsh, D., and Lewtas, J. (1996) DNA adducts and chronic degenerative disease. Pathogenetic relevance and implications in preventive medicine, *Mutat. Res 366*, 197-238.
169. Martin, L. J. (2008) DNA damage and repair: relevance to mechanisms of neurodegeneration, *J. Neuropathol. Exp. Neurol 67*, 377-387.
170. Malik, Q., and Herbert, K. E. (2012) Oxidative and non-oxidative DNA damage and cardiovascular disease, *Free Radic. Res 46*, 554-564.
171. Shukla, P. C., Singh, K. K., Yanagawa, B., Teoh, H., and Verma, S. (2010) DNA damage repair and cardiovascular diseases, *Can. J. Cardiol 26 Suppl A*, 13A-16A.
172. Hecht, S. S. (1999) Tobacco smoke carcinogens and lung cancer, *J. Natl. Cancer Inst 91*, 1194-1210.
173. Lavin, M. F., and Shiloh, Y. (1997) The genetic defect in ataxia-telangiectasia, *Annu. Rev. Immunol 15*, 177-202.

174. DiGiovanna, J. J., and Kraemer, K. H. (2012) Shining a light on xeroderma pigmentosum, *J. Invest Dermatol 132*, 785-796.
175. Lehmann, A. R. (2011) DNA polymerases and repair synthesis in NER in human cells, *DNA Repair (Amst) 10*, 730-733.
176. Lessel, D., Vaz, B., Halder, S., Lockhart, P. J., Marinovic-Terzic, I., Lopez-Mosqueda, J., Philipp, M., Sim, J. C., Smith, K. R., Oehler, J., Cabrera, E., Freire, R., Pope, K., Nahid, A., Norris, F., Leventer, R. J., Delatycki, M. B., Barbi, G., von Ameln, S., Hogel, J., Degoricija, M., Fertig, R., Burkhalter, M. D., Hofmann, K., Thiele, H., Altmuller, J., Nurnberg, G., Nurnberg, P., Bahlo, M., Martin, G. M., Aalfs, C. M., Oshima, J., Terzic, J., Amor, D. J., Dikic, I., Ramadan, K., and Kubisch, C. (2014) Mutations in SPRTN cause early onset hepatocellular carcinoma, genomic instability and progeroid features, *Nat Genet 46*, 1239-1244.
177. Pelz, N., Dempster, N. M., and Shore, P. R. (1990) Analysis of low molecular weight hydrocarbons including 1,3-butadiene in engine exhaust gases using an aluminum oxide porous-layer open-tubular fused-silica column, *J. Chromatogr. Sci 28*, 230-235.
178. Himmelstein, M. W., Acquavella, J. F., Recio, L., Medinsky, M. A., and Bond, J. A. (1997) Toxicology and epidemiology of 1,3-butadiene, *Crit Rev. Toxicol 27*, 1-108.
179. White, W. C. (2007) Butadiene production process overview, *Chem. Biol. Interact 166*, 10-14.
180. Fujita, E. M., Zielinska, B., Campbell, D. E., Sagebiel, J. C., and Ollison, W. (2016) High-end exposure relationships of volatile air toxics and carbon monoxide to community-scale air monitoring stations in Atlanta, Chicago, and Houston, *Air Qual. Atmos. Health 9*, 311-323.
181. Brunnemann, K. D., Kagan, M. R., Cox, J. E., and Hoffmann, D. (1990) Analysis of 1,3-butadiene and other selected gas-phase components in cigarette mainstream and sidestream smoke by gas chromatography-mass selective detection, *Carcinogenesis 11*, 1863-1868.
182. Hoffmann, D., Hoffmann, I., and El Bayoumy, K. (2001) The less harmful cigarette: a controversial issue. a tribute to Ernst L. Wynder, *Chem. Res. Toxicol 14*, 767-790.
183. Clapp, R. W., Jacobs, M. M., and Loechler, E. L. (2008) Environmental and occupational causes of cancer: new evidence 2005-2007, *Rev. Environ. Health 23*, 1-37.
184. Cheng, H., Sathiakumar, N., Graff, J., Matthews, R., and Delzell, E. (2007) 1,3-Butadiene and leukemia among synthetic rubber industry workers: exposure-response relationships, *Chem. Biol. Interact 166*, 15-24.
185. Sielken, R. L., Jr., Valdez-Flores, C., Gargas, M. L., Kirman, C. R., Teta, M. J., and Delzell, E. (2007) Cancer risk assessment for 1,3-butadiene: dose-response modeling from an epidemiological perspective, *Chem. Biol. Interact 166*, 140-149.
186. Graff, J. J., Sathiakumar, N., Macaluso, M., Maldonado, G., Matthews, R., and Delzell, E. (2005) Chemical exposures in the synthetic rubber industry and lymphohematopoietic cancer mortality, *J. Occup. Environ. Med 47*, 916-932.

187. Delzell, E., Sathiakumar, N., Graff, J., and Matthews, R. (2005) Styrene and ischemic heart disease mortality among synthetic rubber industry workers, *J. Occup. Environ. Med 47*, 1235-1243.
188. Sathiakumar, N., Graff, J., Macaluso, M., Maldonado, G., Matthews, R., and Delzell, E. (2005) An updated study of mortality among North American synthetic rubber industry workers, *Occup. Environ. Med 62*, 822-829.
189. Macaluso, M., Larson, R., Delzell, E., Sathiakumar, N., Hovinga, M., Julian, J., Muir, D., and Cole, P. (1996) Leukemia and cumulative exposure to butadiene, styrene and benzene among workers in the synthetic rubber industry, *Toxicology 113*, 190-202.
190. Divine, B. J., and Hartman, C. M. (2001) A cohort mortality study among workers at a 1,3 butadiene facility, *Chem. Biol. Interact 135-136*, 535-553.
191. (2008) **IARC monographs on the evaluation of carcinogenic risks to humans. Volume 97. 1,3-butadiene, ethylene oxide and vinyl halides (vinyl fluoride, vinyl chloride and vinyl bromide)**, *IARC Monogr Eval. Carcinog. Risks Hum 97*, 3-471.
192. Duescher, R. J., and Elfarra, A. A. (1994) Human liver microsomes are efficient catalysts of 1,3-butadiene oxidation: evidence for major roles by cytochromes P450 2A6 and 2E1, *Arch. Biochem. Biophys 311*, 342-349.
193. Csanady, G. A., Guengerich, F. P., and Bond, J. A. (1992) Comparison of the biotransformation of 1,3-butadiene and its metabolite, butadiene monoepoxide, by hepatic and pulmonary tissues from humans, rats and mice, *Carcinogenesis 13*, 1143-1153.
194. Krause, R. J., and Elfarra, A. A. (1997) Oxidation of butadiene monoxide to meso- and (+/-)-diepoxybutane by cDNA-expressed human cytochrome P450s and by mouse, rat, and human liver microsomes: evidence for preferential hydration of meso-diepoxybutane in rat and human liver microsomes, *Arch. Biochem. Biophys 337*, 176-184.
195. Krause, R. J., Sharer, J. E., and Elfarra, A. A. (1997) Epoxide hydrolase-dependent metabolism of butadiene monoxide to 3-butene-1,2-diol in mouse, rat, and human liver, *Drug Metab Dispos 25*, 1013-1015.
196. Carmical, J. R., Zhang, M., Nechev, L., Harris, C. M., Harris, T. M., and Lloyd, R. S. (2000) Mutagenic potential of guanine N2 adducts of butadiene mono- and diolepoxide, *Chem. Res. Toxicol 13*, 18-25.
197. Carmical, J. R., Nechev, L. V., Harris, C. M., Harris, T. M., and Lloyd, R. S. (2000) Mutagenic potential of adenine N(6) adducts of monoepoxide and diolepoxide derivatives of butadiene, *Environ. Mol. Mutagen 35*, 48-56.
198. Van Sittert, N. J., Megens, H. J., Watson, W. P., and Boogaard, P. J. (2000) Biomarkers of exposure to 1,3-butadiene as a basis for cancer risk assessment, *Toxicol. Sci 56*, 189-202.
199. Tretyakova, N. Y., Lin, Y. P., Upton, P. B., Sangaiah, R., and Swenberg, J. A. (1996) Macromolecular adducts of butadiene, *Toxicology 113*, 70-76.
200. Tretyakova, N., Lin, Y., Sangaiah, R., Upton, P. B., and Swenberg, J. A. (1997) Identification and quantitation of DNA adducts from calf thymus DNA exposed to 3,4-epoxy-1-butene, *Carcinogenesis 18*, 137-147.

201. Tretyakova, N. Y., Sangaiah, R., Yen, T. Y., and Swenberg, J. A. (1997) Synthesis, characterization, and in vitro quantitation of N-7-guanine adducts of diepoxybutane, *Chem. Res. Toxicol 10*, 779-785.
202. Sangaraju, D., Villalta, P. W., Wickramaratne, S., Swenberg, J., and Tretyakova, N. (2014) NanoLC/ESI+ HRMS3 quantitation of DNA adducts induced by 1,3-butadiene, *J. Am. Soc. Mass Spectrom 25*, 1124-1135.
203. Sangaraju, D., Goggin, M., Walker, V., Swenberg, J., and Tretyakova, N. (2012) NanoHPLC-nanoESI(+)-MS/MS quantitation of bis-N7-guanine DNA-DNA cross-links in tissues of B6C3F1 mice exposed to subppm levels of 1,3-butadiene, *Anal. Chem 84*, 1732-1739.
204. Citti, L., Gervasi, P. G., Turchi, G., Bellucci, G., and Bianchini, R. (1984) The reaction of 3,4-epoxy-1-butene with deoxyguanosine and DNA in vitro: synthesis and characterization of the main adducts, *Carcinogenesis 5*, 47-52.
205. Fernandes, P. H., Hackfeld, L. C., Kozekov, I. D., Hodge, R. P., and Lloyd, R. S. (2006) Synthesis and mutagenesis of the butadiene-derived N3 2'-deoxyuridine adducts, *Chem Res Toxicol 19*, 968-976.
206. Koc, H., Tretyakova, N. Y., Walker, V. E., Henderson, R. F., and Swenberg, J. A. (1999) Molecular dosimetry of N-7 guanine adduct formation in mice and rats exposed to 1,3-butadiene, *Chem. Res. Toxicol 12*, 566-574.
207. Sangaraju, D., Villalta, P., Goggin, M., Agunsoye, M. O., Campbell, C., and Tretyakova, N. (2013) Capillary HPLC-accurate mass MS/MS quantitation of N7-(2,3,4-trihydroxybut-1-yl)-guanine adducts of 1,3-butadiene in human leukocyte DNA, *Chem. Res. Toxicol 26*, 1486-1497.
208. Tretyakova, N., Sangaiah, R., Yen, T. Y., Gold, A., and Swenberg, J. A. (1997) Adenine adducts with diepoxybutane: isolation and analysis in exposed calf thymus DNA, *Chem Res Toxicol 10*, 1171-1179.
209. Goggin, M., Seneviratne, U., Swenberg, J. A., Walker, V. E., and Tretyakova, N. (2010) Column switching HPLC-ESI(+)-MS/MS methods for quantitative analysis of exocyclic dA adducts in the DNA of laboratory animals exposed to 1,3-butadiene, *Chem. Res. Toxicol 23*, 808-812.
210. Zhang, X. Y., and Elfarra, A. A. (2005) Reaction of 1,2,3,4-diepoxybutane with 2'-deoxyguanosine: initial products and their stabilities and decomposition patterns under physiological conditions, *Chem. Res. Toxicol 18*, 1316-1323.
211. Kligerman, A. D., DeMarini, D. M., Doerr, C. L., Hanley, N. M., Milholland, V. S., and Tennant, A. H. (1999) Comparison of cytogenetic effects of 3,4-epoxy-1-butene and 1,2:3, 4-diepoxybutane in mouse, rat and human lymphocytes following in vitro G0 exposures, *Mutat Res 439*, 13-23.
212. Cochrane, J. E., and Skopek, T. R. (1994) Mutagenicity of butadiene and its epoxide metabolites: I. Mutagenic potential of 1,2-epoxybutene, 1,2,3,4-diepoxybutane and 3,4-epoxy-1,2-butanediol in cultured human lymphoblasts, *Carcinogenesis 15*, 713-717.
213. Verly, W. G., Brakier, L., and Feit, P. W. (1971) Inactivation of the T7 coliphage by the diepoxybutane stereoisomers, *Biochim Biophys Acta 228*, 400-406.
214. Kim, M. Y., Tretyakova, N., and Wogan, G. N. (2007) Mutagenesis of the supF gene by stereoisomers of 1,2,3,4-diepoxybutane, *Chem Res Toxicol 20*, 790-797.

215. Park, S., Anderson, C., Loeber, R., Seetharaman, M., Jones, R., and Tretyakova, N. (2005) Interstrand and intrastrand DNA-DNA cross-linking by 1,2,3,4-diepoxybutane: role of stereochemistry, *J Am Chem Soc 127*, 14355-14365.
216. Millard, J. T., Hanly, T. C., Murphy, K., and Tretyakova, N. (2006) The 5'-GNC site for DNA interstrand cross-linking is conserved for diepoxybutane stereoisomers, *Chem Res Toxicol 19*, 16-19.
217. Goggin, M., Loeber, R., Park, S., Walker, V., Wickliffe, J., and Tretyakova, N. (2007) HPLC-ESI+-MS/MS analysis of N7-guanine-N7-guanine DNA cross-links in tissues of mice exposed to 1,3-butadiene, *Chem Res Toxicol 20*, 839-847.
218. Park, S., Hodge, J., Anderson, C., and Tretyakova, N. (2004) Guanine-adenine DNA cross-linking by 1,2,3,4-diepoxybutane: potential basis for biological activity, *Chem Res Toxicol 17*, 1638-1651.
219. Hsiung, H., and Lown, J. W. (1976) Effects of alkylation by dimethyl sulfate, nitrogen mustard, and mitomycin C on DNA structure as studied by the ethidium binding assay, *Can. J. Biochem 54*, 1047-1054.
220. McGhee, J. D., and Felsenfeld, G. (1979) Reaction of nucleosome DNA with dimethyl sulfate, *Proc. Natl. Acad. Sci. U. S. A 76*, 2133-2137.
221. Craddock, V. M., and Henderson, A. R. (1986) Effect of N-nitrosamines carcinogenic for oesophagus on O6-alkyl-guanine-DNA-methyl transferase in rat oesophagus and liver, *J. Cancer Res. Clin. Oncol 111*, 229-236.
222. Johnston, D. S., and Stone, M. P. (2000) Replication of a site-specific trans-8,9-dihydro-8-(N7-guanyl)-9-hydroxyaflatoxin B(1) adduct by the exonuclease deficient klenow fragment of DNA polymerase I, *Chem. Res. Toxicol 13*, 1158-1164.
223. D'Andrea, A. D., and Haseltine, W. A. (1978) Modification of DNA by aflatoxin B1 creates alkali-labile lesions in DNA at positions of guanine and adenine, *Proc. Natl. Acad. Sci. U. S. A 75*, 4120-4124.
224. Tomasz, M., Lipman, R., Lee, M. S., Verdine, G. L., and Nakanishi, K. (1987) Reaction of acid-activated mitomycin C with calf thymus DNA and model guanines: elucidation of the base-catalyzed degradation of N7-alkylguanine nucleosides, *Biochemistry 26*, 2010-2027.
225. Maccubbin, A. E., Caballes, L., Riordan, J. M., Huang, D. H., and Gurtoo, H. L. (1991) A cyclophosphamide/DNA phosphoester adduct formed in vitro and in vivo, *Cancer Res 51*, 886-892.
226. Hartley, J. A., Hazrati, A., Kelland, L. R., Khanim, R., Shipman, M., Suzenet, F., and Walker, L. F. (2000) A Synthetic Azinomycin Analogue with Demonstrated DNA Cross-Linking Activity: Insights into the Mechanism of Action of this Class of Antitumor Agent The authors gratefully acknowledge the financial support provided by the CRC and the EPSRC. We are indebted to the EPSRC National Mass Spectrometry Centre for performing mass spectral measurements, and the EPSRC Chemical Database Service at Daresbury.1, *Angew. Chem. Int. Ed Engl 39*, 3467-3470.
227. Tudek, B., Boiteux, S., and Laval, J. (1992) Biological properties of imidazole ring-opened N7-methylguanine in M13mp18 phage DNA, *Nucleic Acids Res 20*, 3079-3084.

228. Graziewicz, M. A., Zastawny, T. H., Olinski, R., Speina, E., Siedlecki, J., and Tudek, B. (2000) Fapyadenine is a moderately efficient chain terminator for prokaryotic DNA polymerases, *Free Radic. Biol. Med 28*, 75-83.
229. Giri, I., and Stone, M. P. (2002) Thermal stabilization of the DNA duplex by adducts of aflatoxin B1, *Biopolymers 65*, 190-201.
230. Brown, K. L., Deng, J. Z., Iyer, R. S., Iyer, L. G., Voehler, M. W., Stone, M. P., Harris, C. M., and Harris, T. M. (2006) Unraveling the aflatoxin-FAPY conundrum: structural basis for differential replicative processing of isomeric forms of the formamidopyrimidine-type DNA adduct of aflatoxin B1, *J. Am. Chem. Soc 128*, 15188-15199.
231. Christov, P. P., Brown, K. L., Kozekov, I. D., Stone, M. P., Harris, T. M., and Rizzo, C. J. (2008) Site-specific synthesis and characterization of oligonucleotides containing an N6-(2-deoxy-D-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-N-methylfor mamidopyrimidine lesion, the ring-opened product from N7-methylation of deoxyguanosine, *Chem Res Toxicol 21*, 2324-2333.
232. Christov, P. P., Son, K. J., and Rizzo, C. J. (2014) Synthesis and characterization of oligonucleotides containing a nitrogen mustard formamidopyrimidine monoadduct of deoxyguanosine, *Chem Res Toxicol 27*, 1610-1618.
233. Tudek, B. (2003) Imidazole ring-opened DNA purines and their biological significance, *J. Biochem. Mol. Biol 36*, 12-19.
234. Christov, P. P., Yamanaka, K., Choi, J. Y., Takata, K., Wood, R. D., Guengerich, F. P., Lloyd, R. S., and Rizzo, C. J. (2012) Replication of the 2,6-diamino-4-hydroxy-N(5)-(methyl)-formamidopyrimidine (MeFapy-dGuo) adduct by eukaryotic DNA polymerases, *Chem. Res. Toxicol 25*, 1652-1661.
235. Bailey, E. A., Iyer, R. S., Stone, M. P., Harris, T. M., and Essigmann, J. M. (1996) Mutational properties of the primary aflatoxin B1-DNA adduct, *Proc. Natl. Acad. Sci. U. S. A 93*, 1535-1539.
236. O'Connor, T. R., Boiteux, S., and Laval, J. (1988) Ring-opened 7-methylguanine residues in DNA are a block to in vitro DNA synthesis, *Nucleic Acids Res 16*, 5879-5894.
237. Asagoshi, K., Terato, H., Ohyama, Y., and Ide, H. (2002) Effects of a guanine-derived formamidopyrimidine lesion on DNA replication: translesion DNA synthesis, nucleotide insertion, and extension kinetics, *J. Biol. Chem 277*, 14589-14597.
238. Minko, I. G., Rizzo, C. J., and Lloyd, R. S. (2017) Mutagenic potential of nitrogen mustard-induced formamidopyrimidine DNA adduct: contribution of the non-canonical alpha-anomer, *J Biol Chem*.
239. Baydar, T., Engin, A. B., Girgin, G., Aydin, S., and Sahin, G. (2005) Aflatoxin and ochratoxin in various types of commonly consumed retail ground samples in Ankara, Turkey, *Ann. Agric. Environ. Med 12*, 193-197.
240. McLean, M., and Dutton, M. F. (1995) Cellular interactions and metabolism of aflatoxin: an update, *Pharmacol. Ther 65*, 163-192.
241. Wang, J. S., Huang, T., Su, J., Liang, F., Wei, Z., Liang, Y., Luo, H., Kuang, S. Y., Qian, G. S., Sun, G., He, X., Kensler, T. W., and Groopman, J. D. (2001) Hepatocellular carcinoma and aflatoxin exposure in Zhuqing Village, Fusui

County, People's Republic of China, *Cancer Epidemiol. Biomarkers Prev 10*, 143-146.
242. Hertzog, P. J., Lindsay, S., Jr., and Garner, R. C. (1980) A high pressure liquid chromatography study on the removal of DNA-bound aflatoxin B1 in rat liver and in vitro, *Carcinogenesis 1*, 787-793.
243. Croy, R. G., and Wogan, G. N. (1981) Temporal patterns of covalent DNA adducts in rat liver after single and multiple doses of aflatoxin B1, *Cancer Res 41*, 197-203.
244. Croy, R. G., and Wogan, G. N. (1981) Quantitative comparison of covalent aflatoxin-DNA adducts formed in rat and mouse livers and kidneys, *J. Natl. Cancer Inst 66*, 761-768.
245. Mao, H., Deng, Z., Wang, F., Harris, T. M., and Stone, M. P. (1998) An intercalated and thermally stable FAPY adduct of aflatoxin B1 in a DNA duplex: structural refinement from 1H NMR, *Biochemistry 37*, 4374-4387.
246. Smela, M. E., Hamm, M. L., Henderson, P. T., Harris, C. M., Harris, T. M., and Essigmann, J. M. (2002) The aflatoxin B(1) formamidopyrimidine adduct plays a major role in causing the types of mutations observed in human hepatocellular carcinoma, *Proc. Natl. Acad. Sci. U. S. A 99*, 6655-6660.
247. McMahon, G., Davis, E. F., Huber, L. J., Kim, Y., and Wogan, G. N. (1990) Characterization of c-Ki-ras and N-ras oncogenes in aflatoxin B1-induced rat liver tumors, *Proc. Natl. Acad. Sci. U. S. A 87*, 1104-1108.
248. Lechner, J. F., Cole, K. E., Reddel, R. R., Anderson, L., and Harris, C. C. (1989) Replicative cultures of adult human and rhesus monkey liver epithelial cells, *Cancer Detect. Prev 14*, 239-244.
249. Hsu, I. C., Metcalf, R. A., Sun, T., Welsh, J. A., Wang, N. J., and Harris, C. C. (1991) Mutational hotspot in the p53 gene in human hepatocellular carcinomas, *Nature 350*, 427-428.
250. Bressac, B., Kew, M., Wands, J., and Ozturk, M. (1991) Selective G to T mutations of p53 gene in hepatocellular carcinoma from southern Africa, *Nature 350*, 429-431.
251. Beranek, D. T., Weis, C. C., Evans, F. E., Chetsanga, C. J., and Kadlubar, F. F. (1983) Identification of N5-methyl-N5-formyl-2,5,6-triamino-4-hydroxypyrimidine as a major adduct in rat liver DNA after treatment with the carcinogens, N,N-dimethylnitrosamine or 1,2-dimethylhydrazine, *Biochem. Biophys. Res. Commun 110*, 625-631.
252. Boysen, G., Pachkowski, B. F., Nakamura, J., and Swenberg, J. A. (2009) The formation and biological significance of N7-guanine adducts, *Mutat. Res 678*, 76-94.
253. Graziewicz, M. A., Zastawny, T. H., Olinski, R., and Tudek, B. (1999) SOS-dependent A-->G transitions induced by hydroxyl radical generating system hypoxanthine/xanthine oxidase/Fe3+/EDTA are accompanied by the increase of Fapy-adenine content in M13 mp18 phage DNA, *Mutat. Res 434*, 41-52.
254. Chetsanga, C. J., and Lindahl, T. (1979) Release of 7-methylguanine residues whose imidazole rings have been opened from damaged DNA by a DNA glycosylase from Escherichia coli, *Nucleic Acids Res 6*, 3673-3684.

255. Breimer, L. H. (1984) Enzymatic excision from gamma-irradiated polydeoxyribonucleotides of adenine residues whose imidazole rings have been ruptured, *Nucleic Acids Res 12*, 6359-6367.
256. Tudek, B., van Zeeland, A. A., Kusmierek, J. T., and Laval, J. (1998) Activity of Escherichia coli DNA-glycosylases on DNA damaged by methylating and ethylating agents and influence of 3-substituted adenine derivatives, *Mutat. Res 407*, 169-176.
257. van der Kemp, P. A., Thomas, D., Barbey, R., de Oliveira, R., and Boiteux, S. (1996) Cloning and expression in Escherichia coli of the OGG1 gene of Saccharomyces cerevisiae, which codes for a DNA glycosylase that excises 7,8-dihydro-8-oxoguanine and 2,6-diamino-4-hydroxy-5-N-methylformamidopyrimidine, *Proc. Natl. Acad. Sci. U. S. A 93*, 5197-5202.
258. Karahalil, B., Girard, P. M., Boiteux, S., and Dizdaroglu, M. (1998) Substrate specificity of the Ogg1 protein of Saccharomyces cerevisiae: excision of guanine lesions produced in DNA by ionizing radiation- or hydrogen peroxide/metal ion-generated free radicals, *Nucleic Acids Res 26*, 1228-1233.
259. Hazra, T. K., Izumi, T., Boldogh, I., Imhoff, B., Kow, Y. W., Jaruga, P., Dizdaroglu, M., and Mitra, S. (2002) Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA, *Proc. Natl. Acad. Sci. U. S. A 99*, 3523-3528.
260. Hazra, T. K., Kow, Y. W., Hatahet, Z., Imhoff, B., Boldogh, I., Mokkapati, S. K., Mitra, S., and Izumi, T. (2002) Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions, *J. Biol. Chem 277*, 30417-30420.
261. Alekseyev, Y. O., Hamm, M. L., and Essigmann, J. M. (2004) Aflatoxin B1 formamidopyrimidine adducts are preferentially repaired by the nucleotide excision repair pathway in vivo, *Carcinogenesis 25*, 1045-1051.
262. Tretyakova, N. Y., Groehler, A., and Ji, S. (2015) DNA-Protein cross-links: formation, structural identities, and biological outcomes, *Acc. Chem. Res 48*, 1631-1644.
263. Barker, S., Weinfeld, M., and Murray, D. (2005) DNA-protein crosslinks: their induction, repair, and biological consequences, *Mutat. Res 589*, 111-135.
264. Nakano, T., Miyamoto-Matsubara, M., Shoulkamy, M. I., Salem, A. M., Pack, S. P., Ishimi, Y., and Ide, H. (2013) Translocation and stability of replicative DNA helicases upon encountering DNA-protein cross-links, *J. Biol. Chem 288*, 4649-4658.
265. Yeo, J. E., Wickramaratne, S., Khatwani, S., Wang, Y. C., Vervacke, J., Distefano, M. D., and Tretyakova, N. Y. (2014) Synthesis of site-specific DNA-protein conjugates and their effects on DNA replication, *ACS Chem Biol 9*, 1860-1868.
266. Wickramaratne, S., Boldry, E. J., Buehler, C., Wang, Y. C., Distefano, M. D., and Tretyakova, N. Y. (2015) Error-prone translesion synthesis past DNA-peptide cross-links conjugated to the major groove of DNA via C5 of thymidine, *J. Biol. Chem 290*, 775-787.
267. Wickramaratne, S., Ji, S., Mukherjee, S., Su, Y., Pence, M. G., Lior-Hoffmann, L., Fu, I., Broyde, S., Guengerich, F. P., Distefano, M., Scharer, O. D., Sham, Y. Y., and Tretyakova, N. (2016) Bypass of DNA-Protein Cross-links Conjugated to the

7-Deazaguanine Position of DNA by Translesion Synthesis Polymerases, *J. Biol. Chem*.
268. Nakano, T., Ouchi, R., Kawazoe, J., Pack, S. P., Makino, K., and Ide, H. (2012) T7 RNA polymerases backed up by covalently trapped proteins catalyze highly error prone transcription, *J. Biol. Chem 287*, 6562-6572.
269. Zahn, R. K., Zahn-Daimler, G., Ax, S., Hosokawa, M., and Takeda, T. (1999) Assessment of DNA-protein crosslinks in the course of aging in two mouse strains by use of a modified alkaline filter elution applied to whole tissue samples, *Mech. Ageing Dev 108*, 99-112.
270. Izzotti, A., Cartiglia, C., Taningher, M., De Flora, S., and Balansky, R. (1999) Age-related increases of 8-hydroxy-2'-deoxyguanosine and DNA-protein crosslinks in mouse organs, *Mutat. Res 446*, 215-223.
271. Hubal, E. A., Schlosser, P. M., Conolly, R. B., and Kimbell, J. S. (1997) Comparison of inhaled formaldehyde dosimetry predictions with DNA-protein cross-link measurements in the rat nasal passages, *Toxicol. Appl. Pharmacol 143*, 47-55.
272. Shaham, J., Bomstein, Y., Gurvich, R., Rashkovsky, M., and Kaufman, Z. (2003) DNA-protein crosslinks and p53 protein expression in relation to occupational exposure to formaldehyde, *Occup. Environ. Med 60*, 403-409.
273. Sang, N., Hou, L., Yun, Y., and Li, G. (2009) SO(2) inhalation induces protein oxidation, DNA-protein crosslinks and apoptosis in rat hippocampus, *Ecotoxicol. Environ. Saf 72*, 879-884.
274. Chiu, S. M., Sokany, N. M., Friedman, L. R., and Oleinick, N. L. (1984) Differential processing of ultraviolet or ionizing radiation-induced DNA-protein cross-links in Chinese hamster cells, *Int. J. Radiat. Biol. Relat Stud. Phys. Chem. Med 46*, 681-690.
275. Barker, S., Weinfeld, M., Zheng, J., Li, L., and Murray, D. (2005) Identification of mammalian proteins cross-linked to DNA by ionizing radiation, *J. Biol. Chem 280*, 33826-33838.
276. ALEXANDER, P., and MOROSON, H. (1962) Cross-linking of deoxyribonucleic acid to protein following ultra-violet irradiation different cells, *Nature 194*, 882-883.
277. Hockensmith, J. W., Kubasek, W. L., Evertsz, E. M., Mesner, L. D., and von Hippel, P. H. (1993) Laser cross-linking of proteins to nucleic acids. II. Interactions of the bacteriophage T4 DNA replication polymerase accessory proteins complex with DNA, *J. Biol. Chem 268*, 15721-15730.
278. Moss, T., Dimitrov, S. I., and Houde, D. (1997) UV-laser crosslinking of proteins to DNA, *Methods 11*, 225-234.
279. Taioli, E., Zhitkovich, A., Kinney, P., Udasin, I., Toniolo, P., and Costa, M. (1995) Increased DNA-protein crosslinks in lymphocytes of residents living in chromium-contaminated areas, *Biol. Trace Elem. Res 50*, 175-180.
280. Zhitkovich, A., Lukanova, A., Popov, T., Taioli, E., Cohen, H., Costa, M., and Toniolo, P. (1996) DNA-protein crosslinks in peripheral lymphocytes of individuals exposed to hexavalent chromium compounds, *Biomarkers 1*, 86-93.

281. Zhitkovich, A., Voitkun, V., Kluz, T., and Costa, M. (1998) Utilization of DNA-protein cross-links as a biomarker of chromium exposure, *Environ. Health Perspect 106 Suppl 4*, 969-974.
282. Patierno, S. R., Sugiyama, M., Basilion, J. P., and Costa, M. (1985) Preferential DNA-protein cross-linking by NiCl2 in magnesium-insoluble regions of fractionated Chinese hamster ovary cell chromatin, *Cancer Res 45*, 5787-5794.
283. Liu, H., Xu-Welliver, M., and Pegg, A. E. (2000) The role of human O(6)-alkylguanine-DNA alkyltransferase in promoting 1,2-dibromoethane-induced genotoxicity in Escherichia coli, *Mutat. Res 452*, 1-10.
284. Valadez, J. G., Liu, L., Loktionova, N. A., Pegg, A. E., and Guengerich, F. P. (2004) Activation of bis-electrophiles to mutagenic conjugates by human O^6-alkylguanine-DNA alkyltransferase, *Chem. Res. Toxicol 17*, 972-982.
285. Shaham, J., Bomstein, Y., Meltzer, A., Kaufman, Z., Palma, E., and Ribak, J. (1996) DNA--protein crosslinks, a biomarker of exposure to formaldehyde--in vitro and in vivo studies, *Carcinogenesis 17*, 121-125.
286. Loeber, R. L., Michaelson-Richie, E. D., Codreanu, S. G., Liebler, D. C., Campbell, C. R., and Tretyakova, N. Y. (2009) Proteomic analysis of DNA-protein cross-linking by antitumor nitrogen mustards, *Chem. Res. Toxicol 22*, 1151-1162.
287. Michaelson-Richie, E. D., Ming, X., Codreanu, S. G., Loeber, R. L., Liebler, D. C., Campbell, C., and Tretyakova, N. Y. (2011) Mechlorethamine-induced DNA-protein cross-linking in human fibrosarcoma (HT1080) cells, *J. Proteome. Res 10*, 2785-2796.
288. Groehler, A., Villalta, P. W., Campbell, C., and Tretyakova, N. (2016) Covalent DNA-Protein Cross-Linking by Phosphoramide Mustard and Nornitrogen Mustard in Human Cells, *Chem. Res. Toxicol 29*, 190-202.
289. Ming, X., Groehler, A., Michaelson-Richie, E. D., Villalta, P. W., Campbell, C., and Tretyakova, N. Y. (2017) Mass Spectrometry Based Proteomics Study of Cisplatin-Induced DNA-Protein Cross-Linking in Human Fibrosarcoma (HT1080) Cells, *Chem. Res. Toxicol 30*, 980-995.
290. Markovits, J., Pommier, Y., Kerrigan, D., Covey, J. M., Tilchen, E. J., and Kohn, K. W. (1987) Topoisomerase II-mediated DNA breaks and cytotoxicity in relation to cell proliferation and the cell cycle in NIH 3T3 fibroblasts and L1210 leukemia cells, *Cancer Res 47*, 2050-2055.
291. Pommier, Y., Barcelo, J. M., Rao, V. A., Sordet, O., Jobson, A. G., Thibaut, L., Miao, Z. H., Seiler, J. A., Zhang, H., Marchand, C., Agama, K., Nitiss, J. L., and Redon, C. (2006) Repair of topoisomerase I-mediated DNA damage, *Prog. Nucleic Acid Res. Mol. Biol 81*, 179-229.
292. Gajewski, E., and Dizdaroglu, M. (1990) Hydroxyl radical induced cross-linking of cytosine and tyrosine in nucleohistone, *Biochemistry 29*, 977-980.
293. Xu, X., Muller, J. G., Ye, Y., and Burrows, C. J. (2008) DNA-protein cross-links between guanine and lysine depend on the mechanism of oxidation for formation of C5 vs C8 guanosine adducts, *J. Am. Chem. Soc 130*, 703-709.
294. Johansen, M. E., Muller, J. G., Xu, X., and Burrows, C. J. (2005) Oxidatively induced DNA-protein cross-linking between single-stranded binding protein and oligodeoxynucleotides containing 8-oxo-7,8-dihydro-2'-deoxyguanosine, *Biochemistry 44*, 5660-5671.

295. Brambilla, G., Sciaba, L., Faggin, P., Finollo, R., Bassi, A. M., Ferro, M., and Marinari, U. M. (1985) Methylglyoxal-induced DNA-protein cross-links and cytotoxicity in Chinese hamster ovary cells, *Carcinogenesis 6*, 683-686.
296. Summerfield, F. W., and Tappel, A. L. (1984) Cross-linking of DNA in liver and testes of rats fed 1,3-propanediol, *Chem. Biol. Interact 50*, 87-96.
297. Heck, H. D., Casanova, M., and Starr, T. B. (1990) Formaldehyde toxicity--new understanding, *Crit Rev. Toxicol 20*, 397-426.
298. O'Sullivan, J., Unzeta, M., Healy, J., O'Sullivan, M. I., Davey, G., and Tipton, K. F. (2004) Semicarbazide-sensitive amine oxidases: enzymes with quite a lot to do, *Neurotoxicology 25*, 303-315.
299. Cloos, P. A., Christensen, J., Agger, K., and Helin, K. (2008) Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease, *Genes Dev 22*, 1115-1140.
300. Hou, H., and Yu, H. (2010) Structural insights into histone lysine demethylation, *Curr. Opin. Struct. Biol 20*, 739-748.
301. Swenberg, J. A., Moeller, B. C., Lu, K., Rager, J. E., Fry, R. C., and Starr, T. B. (2013) Formaldehyde carcinogenicity research: 30 years and counting for mode of action, epidemiology, and cancer risk assessment, *Toxicol. Pathol 41*, 181-189.
302. Kawanishi, M., Matsuda, T., and Yagi, T. (2014) Genotoxicity of formaldehyde: molecular basis of DNA damage and mutation, *Frontiers in Environmental Science 2*.
303. Wang, M., Cheng, G., Balbo, S., Carmella, S. G., Villalta, P. W., and Hecht, S. S. (2009) Clear differences in levels of a formaldehyde-DNA adduct in leukocytes of smokers and nonsmokers, *Cancer Res 69*, 7170-7174.
304. Merk, O., and Speit, G. (1998) Significance of formaldehyde-induced DNA-protein crosslinks for mutagenesis, *Environ. Mol. Mutagen 32*, 260-268.
305. Shaham, J., Bomstein, Y., Melzer, A., and Ribak, J. (1997) DNA-Protein Crosslinks and Sister Chromatid Exchanges as Biomarkers of Exposure to Formaldehyde, *Int. J. Occup. Environ. Health 3*, 95-104.
306. Casanova, M., Deyo, D. F., and Heck, H. D. (1989) Covalent binding of inhaled formaldehyde to DNA in the nasal mucosa of Fischer 344 rats: analysis of formaldehyde and DNA by high-performance liquid chromatography and provisional pharmacokinetic interpretation, *Fundam. Appl. Toxicol 12*, 397-417.
307. Casanova, M., Morgan, K. T., Steinhagen, W. H., Everitt, J. I., Popp, J. A., and Heck, H. D. (1991) Covalent binding of inhaled formaldehyde to DNA in the respiratory tract of rhesus monkeys: pharmacokinetics, rat-to-monkey interspecies scaling, and extrapolation to man, *Fundam. Appl. Toxicol 17*, 409-428.
308. Lu, K., Ye, W., Zhou, L., Collins, L. B., Chen, X., Gold, A., Ball, L. M., and Swenberg, J. A. (2010) Structural characterization of formaldehyde-induced cross-links between amino acids and deoxynucleosides and their oligomers, *J. Am. Chem. Soc 132*, 3388-3399.
309. Quievryn, G., and Zhitkovich, A. (2000) Loss of DNA-protein crosslinks from formaldehyde-exposed cells occurs through spontaneous hydrolysis and an active repair process linked to proteosome function, *Carcinogenesis 21*, 1573-1580.

310. Soldi, M., and Bonaldi, T. (2013) The proteomic investigation of chromatin functional domains reveals novel synergisms among distinct heterochromatin components, *Mol. Cell Proteomics 12*, 764-780.
311. Soldi, M., and Bonaldi, T. (2014) The ChroP approach combines ChIP and mass spectrometry to dissect locus-specific proteomic landscapes of chromatin, *J. Vis. Exp*.
312. Wang, C. I., Alekseyenko, A. A., LeRoy, G., Elia, A. E., Gorchakov, A. A., Britton, L. M., Elledge, S. J., Kharchenko, P. V., Garcia, B. A., and Kuroda, M. I. (2013) Chromatin proteins captured by ChIP-mass spectrometry are linked to dosage compensation in Drosophila, *Nat. Struct. Mol. Biol 20*, 202-209.
313. Costa, M., Zhitkovich, A., and Toniolo, P. (1993) DNA-protein cross-links in welders: molecular implications, *Cancer Res 53*, 460-463.
314. Medeiros, M. G., Rodrigues, A. S., Batoreu, M. C., Laires, A., Rueff, J., and Zhitkovich, A. (2003) Elevated levels of DNA-protein crosslinks and micronuclei in peripheral lymphocytes of tannery workers exposed to trivalent chromium, *Mutagenesis 18*, 19-24.
315. Costa, M., Zhitkovich, A., Taioli, E., and Toniolo, P. (1993) Preliminary report on a simple new assay for DNA-protein cross-links as a biomarker of exposures experienced by welders, *J. Toxicol. Environ. Health 40*, 217-222.
316. Toniolo, P., Zhitkovich, A., and Costa, M. (1993) Development and utilization of a new simple assay for DNA-protein crosslinks as a biomarker of exposure to welding fumes, *Int. Arch. Occup. Environ. Health 65*, S87-S89.
317. Kuykendall, J. R., Miller, K. L., Mellinger, K. N., and Cain, A. V. (2006) Waterborne and dietary hexavalent chromium exposure causes DNA-protein crosslink (DPX) formation in erythrocytes of largemouth bass (Micropterus salmoides), *Aquat. Toxicol 78*, 27-31.
318. (2012) **IARC monographs on the evaluation of carcinogenic risks to humans. Volume 100C. (Arsenic, metals, fibres, and dust)**, *IARC Monogr Eval. Carcinog. Risks Hum 100C*, 147 - 167.
319. Alexander, J., and Aaseth, J. (1995) Uptake of chromate in human red blood cells and isolated rat liver cells: the role of the anion carrier, *Analyst 120*, 931-933.
320. Arslan, P., Beltrame, M., and Tomasi, A. (1987) Intracellular chromium reduction, *Biochim. Biophys. Acta 931*, 10-15.
321. Suzuki, Y., and Fukuda, K. (1990) Reduction of hexavalent chromium by ascorbic acid and glutathione with special reference to the rat lung, *Arch. Toxicol 64*, 169-176.
322. Zhitkovich, A., Voitkun, V., and Costa, M. (1995) Glutathione and free amino acids form stable complexes with DNA following exposure of intact mammalian cells to chromate, *Carcinogenesis 16*, 907-913.
323. Mattagajasingh, S. N., and Misra, H. P. (1996) Mechanisms of the carcinogenic chromium(VI)-induced DNA-protein cross-linking and their characterization in cultured intact human cells, *J. Biol. Chem 271*, 33550-33560.
324. Macfie, A., Hagan, E., and Zhitkovich, A. (2010) Mechanism of DNA-protein cross-linking by chromium, *Chem. Res. Toxicol 23*, 341-347.
325. Arakawa, H., Ahmad, R., Naoui, M., and Tajmir-Riahi, H. A. (2000) A comparative study of calf thymus DNA binding to Cr(III) and Cr(VI) ions. Evidence for the

guanine N-7-chromium-phosphate chelate formation, *J. Biol. Chem 275*, 10150-10153.
326. Emadi, A., Jones, R. J., and Brodsky, R. A. (2009) Cyclophosphamide and cancer: golden anniversary, *Nat. Rev. Clin. Oncol 6*, 638-647.
327. Povirk, L. F., and Shuker, D. E. (1994) DNA damage and mutagenesis induced by nitrogen mustards, *Mutat. Res 318*, 205-226.
328. Kaijser, G. P., Korst, A., Beijnen, J. H., Bult, A., and Underberg, W. J. (1993) The analysis of ifosfamide and its metabolites (review), *Anticancer Res 13*, 1311-1324.
329. Ewig, R. A., and Kohn, K. W. (1977) DNA damage and repair in mouse leukemia L1210 cells treated with nitrogen mustard, 1,3-bis(2-chloroethyl)-1-nitrosourea, and other nitrosoureas, *Cancer Res 37*, 2114-2122.
330. Hansson, J., Lewensohn, R., Ringborg, U., and Nilsson, B. (1987) Formation and removal of DNA cross-links induced by melphalan and nitrogen mustard in relation to drug-induced cytotoxicity in human melanoma cells, *Cancer Res 47*, 2631-2637.
331. Loeber, R., Michaelson, E., Fang, Q., Campbell, C., Pegg, A. E., and Tretyakova, N. (2008) Cross-linking of the DNA repair protein O^6-alkylguanine DNA alkyltransferase to DNA in the presence of antitumor nitrogen mustards, *Chem. Res. Toxicol 21*, 787-795.
332. Dasari, S., and Tchounwou, P. B. (2014) Cisplatin in cancer therapy: molecular mechanisms of action, *Eur. J. Pharmacol 740*, 364-378.
333. Kelland, L. (2007) The resurgence of platinum-based cancer chemotherapy, *Nat. Rev. Cancer 7*, 573-584.
334. Chvalova, K., Brabec, V., and Kasparkova, J. (2007) Mechanism of the formation of DNA-protein cross-links by antitumor cisplatin, *Nucleic Acids Res 35*, 1812-1821.
335. Ming, X., and Ming, X. (2011) I. DNA-Protein Cross-Linking by cis-1,1,2,2-Diamminedichloroplatinum (II) (Cisplatin) II. Formation of 8-oxo-dG and oxazolone lesions with p53 derived DNA sequences in the presence of riboflavin.
336. Tretyakova, N. Y., Michaelson-Richie, E. D., Gherezghiher, T. B., Kurtz, J., Ming, X., Wickramaratne, S., Campion, M., Kanugula, S., Pegg, A. E., and Campbell, C. (2013) DNA-reactive protein monoepoxides induce cell death and mutagenesis in mammalian cells, *Biochemistry 52*, 3171-3181.
337. Wickramaratne, S., Mukherjee, S., Villalta, P. W., Scharer, O. D., and Tretyakova, N. Y. (2013) Synthesis of sequence-specific DNA-protein conjugates via a reductive amination strategy, *Bioconjug. Chem 24*, 1496-1506.
338. Circu, M. L., and Aw, T. Y. (2010) Reactive oxygen species, cellular redox systems, and apoptosis, *Free Radic. Biol. Med 48*, 749-762.
339. Jezek, P., and Hlavata, L. (2005) Mitochondria in homeostasis of reactive oxygen species in cell, tissues, and organism, *Int. J. Biochem. Cell Biol 37*, 2478-2503.
340. Nohl, H., Kozlov, A. V., Gille, L., and Staniek, K. (2003) Cell respiration and formation of reactive oxygen species: facts and artefacts, *Biochem. Soc. Trans 31*, 1308-1311.
341. Forman, H. J., and Torres, M. (2002) Reactive oxygen species and cell signaling: respiratory burst in macrophage signaling, *Am. J. Respir. Crit Care Med 166*, S4-S8.

342. Torres, M. A., Jones, J. D., and Dangl, J. L. (2006) Reactive oxygen species signaling in response to pathogens, *Plant Physiol 141*, 373-378.
343. Valavanidis, A., Vlachogianni, T., and Fiotakis, K. (2009) Tobacco smoke: involvement of reactive oxygen species and stable free radicals in mechanisms of oxidative damage, carcinogenesis and synergistic effects with other respirable particles, *Int. J. Environ. Res. Public Health 6*, 445-462.
344. Church, D. F., and Pryor, W. A. (1985) Free-radical chemistry of cigarette smoke and its toxicological implications, *Environ. Health Perspect 64*, 111-126.
345. Montoliu, C., Valles, S., Renau-Piqueras, J., and Guerri, C. (1994) Ethanol-induced oxygen radical formation and lipid peroxidation in rat brain: effect of chronic alcohol consumption, *J. Neurochem 63*, 1855-1862.
346. Bailey, S. M., Pietsch, E. C., and Cunningham, C. C. (1999) Ethanol stimulates the production of reactive oxygen species at mitochondrial complexes I and III, *Free Radic. Biol. Med 27*, 891-900.
347. O'Brien, T. J., Ceryak, S., and Patierno, S. R. (2003) Complexities of chromium carcinogenesis: role of cellular response, repair and recovery mechanisms, *Mutat. Res 533*, 3-36.
348. Spooner, R., and Yilmaz, O. (2011) The Role of Reactive-Oxygen-Species in Microbial Persistence and Inflammation, *International Journal of Molecular Sciences 12*, 334-352.
349. Lee, S. R., Yang, K. S., Kwon, J., Lee, C., Jeong, W., and Rhee, S. G. (2002) Reversible inactivation of the tumor suppressor PTEN by H2O2, *J. Biol. Chem 277*, 20336-20342.
350. Kwon, J., Lee, S. R., Yang, K. S., Ahn, Y., Kim, Y. J., Stadtman, E. R., and Rhee, S. G. (2004) Reversible oxidation and inactivation of the tumor suppressor PTEN in cells stimulated with peptide growth factors, *Proc. Natl. Acad. Sci. U. S. A 101*, 16419-16424.
351. Seo, J. H., Ahn, Y., Lee, S. R., Yeol, Y. C., and Chung, H. K. (2005) The major target of the endogenously generated reactive oxygen species in response to insulin stimulation is phosphatase and tensin homolog and not phosphoinositide-3 kinase (PI-3 kinase) in the PI-3 kinase/Akt pathway, *Mol. Biol. Cell 16*, 348-357.
352. Kamata, H., Honda, S., Maeda, S., Chang, L., Hirata, H., and Karin, M. (2005) Reactive oxygen species promote TNFalpha-induced death and sustained JNK activation by inhibiting MAP kinase phosphatases, *Cell 120*, 649-661.
353. Robinson, K. A., Stewart, C. A., Pye, Q. N., Nguyen, X., Kenney, L., Salzman, S., Floyd, R. A., and Hensley, K. (1999) Redox-sensitive protein phosphatase activity regulates the phosphorylation state of p38 protein kinase in primary astrocyte culture, *J. Neurosci. Res 55*, 724-732.
354. Ray, P. D., Huang, B. W., and Tsuji, Y. (2012) Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling, *Cell Signal 24*, 981-990.
355. Morgan, M. J., and Liu, Z. G. (2011) Crosstalk of reactive oxygen species and NF-kappaB signaling, *Cell Res 21*, 103-115.
356. Oleinick, N. L., Chiu, S. M., Ramakrishnan, N., and Xue, L. Y. (1987) The formation, identification, and significance of DNA-protein cross-links in mammalian cells, *Br J Cancer Suppl 8*, 135-140.

357. Nakano, T., Terato, H., Asagoshi, K., Masaoka, A., Mukuta, M., Ohyama, Y., Suzuki, T., Makino, K., and Ide, H. (2003) DNA-protein cross-link formation mediated by oxanine. A novel genotoxic mechanism of nitric oxide-induced DNA damage, *J Biol Chem 278*, 25264-25272.
358. Nakano, T., Terato, H., Asagoshi, K., Masaoka, A., Mukuta, M., Ohyama, Y., Suzuki, T., Makino, K., and Ide, H. (2003) DNA-Protein Cross-link Formation Mediated by Oxanine: a novel genotoxi mechanism of nitric oxide-induced DNA damage, *J. Biol. Chem. 278*, 25264-25272.
359. Kroeger, K. M., Hashimoto, M., Kow, Y. W., and Greenberg, M. M. (2003) Cross-linking of 2-deoxyribonolactone and its beta-elimination product by base excision repair enzymes, *Biochemistry 42*, 2449-2455.
360. Guan, L., and Greenberg, M. M. (2010) Irreversible inhibition of DNA polymerase beta by an oxidized abasic lesion, *J. Am. Chem. Soc 132*, 5004-5005.
361. Perrier, S., Hau, J., Gasparutto, D., Cadet, J., Favier, A., and Ravanat, J. L. (2006) Characterization of lysine-guanine cross-links upon one-electron oxidation of a guanine-containing oligonucleotide in the presence of a trilysine peptide, *J. Am. Chem. Soc 128*, 5703-5710.
362. Dizdaroglu, M., Gajewski, E., Reddy, P., and Margolis, S. A. (1989) Structure of a hydroxyl radical induced DNA-protein cross-link involving thymine and tyrosine in nucleohistone, *Biochemistry 28*, 3625-3628.
363. Gajewski, E., Fuciarelli, A. F., and Dizdaroglu, M. (1988) Structure of hydroxyl radical-induced DNA-protein crosslinks in calf thymus nucleohistone in vitro, *Int. J. Radiat. Biol 54*, 445-459.
364. Mozaffarian, D., Benjamin, E. J., Go, A. S., Arnett, D. K., Blaha, M. J., Cushman, M., Das, S. R., de Ferranti, S., Despres, J. P., Fullerton, H. J., Howard, V. J., Huffman, M. D., Isasi, C. R., Jimenez, M. C., Judd, S. E., Kissela, B. M., Lichtman, J. H., Lisabeth, L. D., Liu, S., Mackey, R. H., Magid, D. J., McGuire, D. K., Mohler, E. R., III, Moy, C. S., Muntner, P., Mussolino, M. E., Nasir, K., Neumar, R. W., Nichol, G., Palaniappan, L., Pandey, D. K., Reeves, M. J., Rodriguez, C. J., Rosamond, W., Sorlie, P. D., Stein, J., Towfighi, A., Turan, T. N., Virani, S. S., Woo, D., Yeh, R. W., and Turner, M. B. (2016) Executive Summary: Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association, *Circulation 133*, 447-454.
365. Mozaffarian, D., Benjamin, E. J., Go, A. S., Arnett, D. K., Blaha, M. J., Cushman, M., Das, S. R., de Ferranti, S., Despres, J. P., Fullerton, H. J., Howard, V. J., Huffman, M. D., Isasi, C. R., Jimenez, M. C., Judd, S. E., Kissela, B. M., Lichtman, J. H., Lisabeth, L. D., Liu, S., Mackey, R. H., Magid, D. J., McGuire, D. K., Mohler, E. R., III, Moy, C. S., Muntner, P., Mussolino, M. E., Nasir, K., Neumar, R. W., Nichol, G., Palaniappan, L., Pandey, D. K., Reeves, M. J., Rodriguez, C. J., Rosamond, W., Sorlie, P. D., Stein, J., Towfighi, A., Turan, T. N., Virani, S. S., Woo, D., Yeh, R. W., and Turner, M. B. (2016) Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association, *Circulation 133*, e38-e360.
366. Hashmi, S., and Al Salam, S. (2015) Acute myocardial infarction and myocardial ischemia-reperfusion injury: a comparison, *Int. J. Clin. Exp. Pathol 8*, 8786-8796.

367. Hausenloy, D. J., and Yellon, D. M. (2013) Myocardial ischemia-reperfusion injury: a neglected therapeutic target, *J. Clin. Invest 123*, 92-100.
368. Arroyo, C. M., Kramer, J. H., Dickens, B. F., and Weglicki, W. B. (1987) Identification of free radicals in myocardial ischemia/reperfusion by spin trapping with nitrone DMPO, *FEBS Lett 221*, 101-104.
369. Avkiran, M., and Marber, M. S. (2002) Na(+)/H(+) exchange inhibitors for cardioprotective therapy: progress, problems and prospects, *J. Am. Coll. Cardiol 39*, 747-753.
370. Carbone, F., Nencioni, A., Mach, F., Vuilleumier, N., and Montecucco, F. (2013) Pathophysiological role of neutrophils in acute myocardial infarction, *Thromb. Haemost 110*, 501-514.
371. Duilio, C., Ambrosio, G., Kuppusamy, P., DiPaula, A., Becker, L. C., and Zweier, J. L. (2001) Neutrophils are primary source of O2 radicals during reperfusion after prolonged myocardial ischemia, *Am. J. Physiol Heart Circ. Physiol 280*, H2649-H2657.
372. Lefer, D. J., Shandelya, S. M., Serrano, C. V., Jr., Becker, L. C., Kuppusamy, P., and Zweier, J. L. (1993) Cardioprotective actions of a monoclonal antibody against CD-18 in myocardial ischemia-reperfusion injury, *Circulation 88*, 1779-1787.
373. Morgenfeld, M. C., Pavlovsky, A., Suarez, A., Somoza, N., Pavlovsky, S., Palau, M., and Barros, C. A. (1975) Combined cyclophosphamide vincristine, procarbazine, and prednisone (COPP) therapy of malignant lymphoma. Evaluation of 190 patients, *Cancer 36*, 1241-1249.
374. Cunningham, D., Hawkes, E. A., Jack, A., Qian, W., Smith, P., Mouncey, P., Pocock, C., Ardeshna, K. M., Radford, J. A., McMillan, A., Davies, J., Turner, D., Kruger, A., Johnson, P., Gambell, J., and Linch, D. (2013) Rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisolone in patients with newly diagnosed diffuse large B-cell non-Hodgkin lymphoma: a phase 3 comparison of dose intensification with 14-day versus 21-day cycles, *Lancet 381*, 1817-1826.
375. Moignet, A., Hasanali, Z., Zambello, R., Pavan, L., Bareau, B., Tournilhac, O., Roussel, M., Fest, T., Awwad, A., Baab, K., Semenzato, G., Houot, R., Loughran, T. P., Jr., and Lamy, T. (2014) Cyclophosphamide as a first-line therapy in LGL leukemia, *Leukemia 28*, 1134-1136.
376. Morgan, G. J., Davies, F. E., Gregory, W. M., Russell, N. H., Bell, S. E., Szubert, A. J., Navarro, C. N., Cook, G., Feyler, S., Byrne, J. L., Roddie, H., Rudin, C., Drayson, M. T., Owen, R. G., Ross, F. M., Jackson, G. H., and Child, J. A. (2011) Cyclophosphamide, thalidomide, and dexamethasone (CTD) as initial therapy for patients with multiple myeloma unsuitable for autologous transplantation, *Blood 118*, 1231-1238.
377. Morgan, G. J., Davies, F. E., Gregory, W. M., Bell, S. E., Szubert, A. J., Navarro, C. N., Cook, G., Feyler, S., Johnson, P. R., Rudin, C., Drayson, M. T., Owen, R. G., Ross, F. M., Russell, N. H., Jackson, G. H., and Child, J. A. (2012) Cyclophosphamide, thalidomide, and dexamethasone as induction therapy for newly diagnosed multiple myeloma patients destined for autologous stem-cell

transplantation: MRC Myeloma IX randomized trial results, *Haematologica 97*, 442-450.
378. Yule, S. M., Foreman, N. K., Mitchell, C., Gouldon, N., May, P., and McDowell, H. P. (1997) High-dose cyclophosphamide for poor-prognosis and recurrent pediatric brain tumors: a dose-escalation study, *J. Clin. Oncol 15*, 3258-3265.
379. Al Homsi, A. S., Roy, T. S., Cole, K., Feng, Y., and Duffner, U. (2015) Post-transplant high-dose cyclophosphamide for the prevention of graft-versus-host disease, *Biol. Blood Marrow Transplant 21*, 604-611.
380. Luznik, L., and Fuchs, E. J. (2010) High-dose, post-transplantation cyclophosphamide to promote graft-host tolerance after allogeneic hematopoietic stem cell transplantation, *Immunol. Res 47*, 65-77.
381. Chang, T. K., Weber, G. F., Crespi, C. L., and Waxman, D. J. (1993) Differential activation of cyclophosphamide and ifosphamide by cytochromes P-450 2B and 3A in human liver microsomes, *Cancer Res 53*, 5629-5637.
382. Huang, Z., Roy, P., and Waxman, D. J. (2000) Role of human liver microsomal CYP3A4 and CYP2B6 in catalyzing N-dechloroethylation of cyclophosphamide and ifosfamide, *Biochem. Pharmacol 59*, 961-972.
383. Hemminki, K. (1985) Binding of metabolites of cyclophosphamide to DNA in a rat liver microsomal system and in vivo in mice, *Cancer Res 45*, 4237-4243.
384. Hemminki, K. (1987) DNA-binding products of nornitrogen mustard, a metabolite of cyclophosphamide, *Chem. Biol. Interact 61*, 75-88.
385. Benson, A. J., Martin, C. N., and Garner, R. C. (1988) N-(2-hydroxyethyl)-N-[2-(7-guaninyl)ethyl]amine, the putative major DNA adduct of cyclophosphamide in vitro and in vivo in the rat, *Biochem. Pharmacol 37*, 2979-2985.
386. Bjorklund, C. C., and Davis, W. B. (2007) Stable DNA-protein cross-links are products of DNA charge transport in a nucleosome core particle, *Biochemistry 46*, 10745-10755.
387. Fornace, A. J., Jr., and Little, J. B. (1977) DNA crosslinking induced by x-rays and chemical agents, *Biochim. Biophys. Acta 477*, 343-355.
388. Loeber, R., Rajesh, M., Fang, Q., Pegg, A. E., and Tretyakova, N. (2006) Cross-linking of the human DNA repair protein O^6-alkylguanine DNA alkyltransferase to DNA in the presence of 1,2,3,4-diepoxybutane, *Chem. Res. Toxicol 19*, 645-654.
389. Kohn, K. W., Erickson, L. C., Ewig, R. A., and Friedman, C. A. (1976) Fractionation of DNA from mammalian cells by alkaline elution, *Biochemistry 15*, 4629-4637.
390. Ewig, R. A., and Kohn, K. W. (1978) DNA-protein cross-linking and DNA interstrand cross-linking by haloethylnitrosoureas in L1210 cells, *Cancer Res 38*, 3197-3203.
391. Merk, O., and Speit, G. (1999) Detection of crosslinks with the comet assay in relationship to genotoxicity and cytotoxicity, *Environ. Mol. Mutagen 33*, 167-172.
392. Speit, G., Schutz, P., and Merk, O. (2000) Induction and repair of formaldehyde-induced DNA-protein crosslinks in repair-deficient human cell lines, *Mutagenesis 15*, 85-90.

393. Wang, T. S., Hsu, T. Y., Chung, C. H., Wang, A. S., Bau, D. T., and Jan, K. Y. (2001) Arsenite induces oxidative DNA adducts and DNA-protein cross-links in mammalian cells, *Free Radic. Biol. Med 31*, 321-330.
394. Wozniak, K., and Blasiak, J. (2002) Free radicals-mediated induction of oxidized DNA bases and DNA-protein cross-links by nickel chloride, *Mutat. Res 514*, 233-243.
395. Wozniak, K., and Blasiak, J. (2003) In vitro genotoxicity of lead acetate: induction of single and double DNA strand breaks and DNA-protein cross-links, *Mutat. Res 535*, 127-139.
396. Zhitkovich, A., and Costa, M. (1992) A simple, sensitive assay to detect DNA-protein crosslinks in intact cells and in vivo, *Carcinogenesis 13*, 1485-1489.
397. Nelson, C. A. (1971) The binding of detergents to proteins. I. The maximum amount of dodecyl sulfate bound to proteins and the resistance to binding of several proteins, *J. Biol. Chem 246*, 3895-3901.
398. Sambrook, J., and Russell, D. W. (2006) Purification of nucleic acids by extraction with phenol:chloroform, *CSH. Protoc 2006*.
399. Michaelson-Richie, E. D., Loeber, R. L., Codreanu, S. G., Ming, X., Liebler, D. C., Campbell, C., and Tretyakova, N. Y. (2010) DNA-protein cross-linking by 1,2,3,4-diepoxybutane, *J. Proteome. Res 9*, 4356-4367.
400. Barker, S., Murray, D., Zheng, J., Li, L., and Weinfeld, M. (2005) A method for the isolation of covalent DNA-protein crosslinks suitable for proteomics analysis, *Anal. Biochem 344*, 204-215.
401. Qin, H., and Wang, Y. (2009) Exploring DNA-binding proteins with in vivo chemical cross-linking and mass spectrometry, *J. Proteome. Res 8*, 1983-1991.
402. Chen, H. J., Chiu, W. L., Lin, W. P., and Yang, S. S. (2008) Investigation of DNA-protein cross-link formation between lysozyme and oxanine by mass spectrometry, *Chembiochem 9*, 1074-1081.
403. Doneanu, C. E., Gafken, P. R., Bennett, S. E., and Barofsky, D. F. (2004) Mass spectrometry of UV-cross-linked protein-nucleic acid complexes: identification of amino acid residues in the single-stranded DNA-binding domain of human replication protein A, *Anal. Chem 76*, 5667-5676.
404. Zhang, Y., Fonslow, B. R., Shan, B., Baek, M. C., and Yates, J. R., III. (2013) Protein analysis by shotgun/bottom-up proteomics, *Chem. Rev 113*, 2343-2394.
405. Olsen, J. V., Ong, S. E., and Mann, M. (2004) Trypsin cleaves exclusively C-terminal to arginine and lysine residues, *Mol. Cell Proteomics 3*, 608-614.
406. Gillet, L. C., Leitner, A., and Aebersold, R. (2016) Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing, *Annu. Rev. Anal. Chem. (Palo. Alto. Calif.) 9*, 449-472.
407. Cress, A. E., Kurath, K. M., Stea, B., and Bowden, G. T. (1990) The crosslinking of nuclear protein to DNA using ionizing radiation, *J. Cancer Res. Clin. Oncol 116*, 324-330.
408. Stochaj, W. R., Berkelman, T., and Laird, N. (2007) Mass spectrometry-compatible silver staining, *CSH. Protoc 2007*, db.
409. Wisniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample preparation method for proteome analysis, *Nat. Methods 6*, 359-362.

410. Huang, d. W., Sherman, B. T., and Lempicki, R. A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists, *Nucleic Acids Res 37*, 1-13.
411. Kramer, A., Green, J., Pollard, J., Jr., and Tugendreich, S. (2014) Causal analysis approaches in Ingenuity Pathway Analysis, *Bioinformatics 30*, 523-530.
412. Rosenstein, B. S., Subramanian, D., and Muller, M. T. (1997) The involvement of topoisomerase I in the induction of DNA-protein crosslinks and DNA single-strand breaks in cells of ultraviolet-irradiated human and frog cell lines, *Radiat. Res 148*, 575-579.
413. Kingma, P. S., and Osheroff, N. (1997) Apurinic sites are position-specific topoisomerase II poisons, *J. Biol. Chem 272*, 1148-1155.
414. Bantscheff, M., Lemeer, S., Savitski, M. M., and Kuster, B. (2012) Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present, *Anal. Bioanal. Chem 404*, 939-965.
415. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007) Quantitative mass spectrometry in proteomics: a critical review, *Anal. Bioanal. Chem 389*, 1017-1031.
416. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics, *Mol. Cell Proteomics 1*, 376-386.
417. Byrum, S. D., Raman, A., Taverna, S. D., and Tackett, A. J. (2012) ChAP-MS: a method for identification of proteins and histone posttranslational modifications at a single genomic locus, *Cell Rep 2*, 198-205.
418. Stewart, M. D., Li, J., and Wong, J. (2005) Relationship between histone H3 lysine 9 methylation, transcription repression, and heterochromatin protein 1 recruitment, *Mol. Cell Biol 25*, 2525-2538.
419. Eissenberg, J. C., and Shilatifard, A. (2010) Histone H3 lysine 4 (H3K4) methylation in development and differentiation, *Dev. Biol 339*, 240-249.
420. Pommier, Y., Huang, S. Y., Gao, R., Das, B. B., Murai, J., and Marchand, C. (2014) Tyrosyl-DNA-phosphodiesterases (TDP1 and TDP2), *DNA Repair (Amst) 19*, 114-129.
421. Stingele, J., Schwarz, M. S., Bloemeke, N., Wolf, P. G., and Jentsch, S. (2014) A DNA-dependent protease involved in DNA-protein crosslink repair, *Cell 158*, 327-338.
422. Regairaz, M., Zhang, Y. W., Fu, H., Agama, K. K., Tata, N., Agrawal, S., Aladjem, M. I., and Pommier, Y. (2011) Mus81-mediated DNA cleavage resolves replication forks stalled by topoisomerase I-DNA complexes, *J. Cell Biol 195*, 739-749.
423. Quievryn, G., Messer, J., and Zhitkovich, A. (2002) Carcinogenic chromium(VI) induces cross-linking of vitamin C to DNA in vitro and in human lung A549 cells, *Biochemistry 41*, 3156-3167.
424. Nakano, T., Katafuchi, A., Matsubara, M., Terato, H., Tsuboi, T., Masuda, T., Tatsumoto, T., Pack, S. P., Makino, K., Croteau, D. L., Van Houten, B., Iijima, K., Tauchi, H., and Ide, H. (2009) Homologous recombination but not nucleotide

excision repair plays a pivotal role in tolerance of DNA-protein cross-links in mammalian cells, *J. Biol. Chem 284*, 27065-27076.
425. Nakano, T., Morishita, S., Katafuchi, A., Matsubara, M., Horikawa, Y., Terato, H., Salem, A. M., Izumi, S., Pack, S. P., Makino, K., and Ide, H. (2007) Nucleotide excision repair and homologous recombination systems commit differentially to the repair of DNA-protein crosslinks, *Mol. Cell 28*, 147-158.
426. Reardon, J. T., Cheng, Y., and Sancar, A. (2006) Repair of DNA-protein cross-links in mammalian cells, *Cell Cycle 5*, 1366-1370.
427. Baker, D. J., Wuenschell, G., Xia, L., Termini, J., Bates, S. E., Riggs, A. D., and O'Connor, T. R. (2007) Nucleotide excision repair eliminates unique DNA-protein cross-links from mammalian cells, *J. Biol. Chem 282*, 22592-22604.
428. Stingele, J., Habermann, B., and Jentsch, S. (2015) DNA-protein crosslink repair: proteases as DNA repair enzymes, *Trends Biochem Sci 40*, 67-71.
429. Maskey, R. S., Kim, M. S., Baker, D. J., Childs, B., Malureanu, L. A., Jeganathan, K. B., Machida, Y., van Deursen, J. M., and Machida, Y. J. (2014) Spartan deficiency causes genomic instability and progeroid phenotypes, *Nat. Commun 5*, 5744.
430. Duxin, J. P., Dewar, J. M., Yardimci, H., and Walter, J. C. (2014) Repair of a DNA-protein crosslink by replication-coupled proteolysis, *Cell 159*, 346-357.
431. Stingele, J., Bellelli, R., Alte, F., Hewitt, G., Sarek, G., Maslen, S. L., Tsutakawa, S. E., Borg, A., Kjaer, S., Tainer, J. A., Skehel, J. M., Groll, M., and Boulton, S. J. (2016) Mechanism and Regulation of DNA-Protein Crosslink Repair by the DNA-Dependent Metalloprotease SPRTN, *Mol Cell 64*, 688-703.
432. Mosbech, A., Gibbs-Seymour, I., Kagias, K., Thorslund, T., Beli, P., Povlsen, L., Nielsen, S. V., Smedegaard, S., Sedgwick, G., Lukas, C., Hartmann-Petersen, R., Lukas, J., Choudhary, C., Pocock, R., Bekker-Jensen, S., and Mailand, N. (2012) DVC1 (C1orf124) is a DNA damage-targeting p97 adaptor that promotes ubiquitin-dependent responses to replication blocks, *Nat Struct Mol Biol 19*, 1084-1092.
433. Morocz, M., Zsigmond, E., Toth, R., Enyedi, M. Z., Pinter, L., and Haracska, L. (2017) DNA-dependent protease activity of human Spartan facilitates replication of DNA-protein crosslink-containing DNA, *Nucleic Acids Res 45*, 3172-3188.
434. Chakrabarti, S. K., Bai, C., and Subramanian, K. S. (1999) DNA-Protein crosslinks induced by nickel compounds in isolated rat renal cortical cells and its antagonism by specific amino acids and magnesium ion, *Toxicol. Appl. Pharmacol 154*, 245-255.
435. Baker, J. M., Parish, J. H., and Curtis, J. P. (1984) DNA-DNA and DNA-protein crosslinking and repair in Neurospora crassa following exposure to nitrogen mustard, *Mutat. Res 132*, 171-179.
436. Zwelling, L. A., Anderson, T., and Kohn, K. W. (1979) DNA-protein and DNA interstrand cross-linking by cis- and trans-platinum(II) diamminedichloride in L1210 mouse leukemia cells and relation to cytotoxicity, *Cancer Res 39*, 365-369.
437. Kloster, M., Kostrhunova, H., Zaludova, R., Malina, J., Kasparkova, J., Brabec, V., and Farrell, N. (2004) Trifunctional dinuclear platinum complexes as DNA-protein cross-linking agents, *Biochemistry 43*, 7776-7786.

438. Miller, C. A., III, Cohen, M. D., and Costa, M. (1991) Complexing of actin and other nuclear proteins to DNA by cis-diamminedichloroplatinum(II) and chromium compounds, *Carcinogenesis 12*, 269-276.
439. Brodsky, R. A. (2002) High dose cyclophosphamide treatment for autoimmune disorders, *ScientificWorldJournal 2*, 1808-1815.
440. Gherezghiher, T. B., Ming, X., Villalta, P. W., Campbell, C., and Tretyakova, N. Y. (2013) 1,2,3,4-Diepoxybutane-induced DNA-protein cross-linking in human fibrosarcoma (HT1080) cells, *J. Proteome. Res 12*, 2151-2164.
441. O'Brien, J., Wilson, I., Orton, T., and Pognan, F. (2000) Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity, *Eur. J. Biochem 267*, 5421-5426.
442. Yates, J. R., III, Eng, J. K., and McCormack, A. L. (1995) Mining genomes: correlating tandem mass spectra of modified and unmodified peptides to sequences in nucleotide databases, *Anal. Chem 67*, 3202-3210.
443. Camon, E., Barrell, D., Brooksbank, C., Magrane, M., and Apweiler, R. (2003) The Gene Ontology Annotation (GOA) Project--Application of GO in SWISS-PROT, TrEMBL and InterPro, *Comp Funct. Genomics 4*, 71-74.
444. Schonthaler, H. B., Guinea-Viniegra, J., Wculek, S. K., Ruppen, I., Ximenez-Embun, P., Guio-Carrion, A., Navarro, R., Hogg, N., Ashman, K., and Wagner, E. F. (2013) S100A8-S100A9 protein complex mediates psoriasis by regulating the expression of complement factor C3, *Immunity 39*, 1171-1181.
445. Hublitz, P., Kunowska, N., Mayer, U. P., Muller, J. M., Heyne, K., Yin, N., Fritzsche, C., Poli, C., Miguet, L., Schupp, I. W., van Grunsven, L. A., Potiers, N., van Dorsselaer, A., Metzger, E., Roemer, K., and Schule, R. (2005) NIR is a novel INHAT repressor that modulates the transcriptional activity of p53, *Genes Dev 19*, 2912-2924.
446. Diaz, V. M., Mori, S., Longobardi, E., Menendez, G., Ferrai, C., Keough, R. A., Bachi, A., and Blasi, F. (2007) p160 Myb-binding protein interacts with Prep1 and inhibits its transcriptional activity, *Mol. Cell Biol 27*, 7981-7990.
447. Mori, S., Bernardi, R., Laurent, A., Resnati, M., Crippa, A., Gabrieli, A., Keough, R., Gonda, T. J., and Blasi, F. (2012) Myb-binding protein 1A (MYBBP1A) is essential for early embryonic development, controls cell cycle and mitosis, and acts as a tumor suppressor, *PLoS. One 7*, e39723.
448. McKay, M. J., Troelstra, C., van der, S. P., Kanaar, R., Smit, B., Hagemeijer, A., Bootsma, D., and Hoeijmakers, J. H. (1996) Sequence conservation of the rad21 Schizosaccharomyces pombe DNA double-strand break repair gene in human and mouse, *Genomics 36*, 305-315.
449. Pan, M. R., Hsieh, H. J., Dai, H., Hung, W. C., Li, K., Peng, G., and Lin, S. Y. (2012) Chromodomain helicase DNA-binding protein 4 (CHD4) regulates homologous recombination DNA repair, and its deficiency sensitizes cells to poly(ADP-ribose) polymerase (PARP) inhibitor treatment, *J. Biol. Chem 287*, 6764-6772.
450. Argraves, W. S., Suzuki, S., Arai, H., Thompson, K., Pierschbacher, M. D., and Ruoslahti, E. (1987) Amino acid sequence of the human fibronectin receptor, *J. Cell Biol 105*, 1183-1190.

451. Soker, S., Takashima, S., Miao, H. Q., Neufeld, G., and Klagsbrun, M. (1998) Neuropilin-1 is expressed by endothelial and tumor cells as an isoform-specific receptor for vascular endothelial growth factor, *Cell 92*, 735-745.
452. Fu, G. K., and Markovitz, D. M. (1998) The human LON protease binds to mitochondrial promoters in a single-stranded, site-specific, strand-specific manner, *Biochemistry 37*, 1905-1909.
453. Gabrielli, F., Donadel, G., Bensi, G., Heguy, A., and Melli, M. (1995) A nuclear protein, synthesized in growth-arrested human hepatoblastoma cells, is a novel member of the short-chain alcohol dehydrogenase family, *Eur. J. Biochem 232*, 473-477.
454. Vandekerckhove, J., and Weber, K. (1978) Mammalian cytoplasmic actins are the products of at least two genes and differ in primary structure in at least 25 identified positions from skeletal muscle actins, *Proc. Natl. Acad. Sci. U. S. A 75*, 1106-1110.
455. Orphanides, G., LeRoy, G., Chang, C. H., Luse, D. S., and Reinberg, D. (1998) FACT, a factor that facilitates transcript elongation through nucleosomes, *Cell 92*, 105-116.
456. Orphanides, G., Wu, W. H., Lane, W. S., Hampsey, M., and Reinberg, D. (1999) The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins, *Nature 400*, 284-288.
457. Wada, T., Orphanides, G., Hasegawa, J., Kim, D. K., Shima, D., Yamaguchi, Y., Fukuda, A., Hisatake, K., Oh, S., Reinberg, D., and Handa, H. (2000) FACT relieves DSIF/NELF-mediated inhibition of transcriptional elongation and reveals functional differences between P-TEFb and TFIIH, *Mol. Cell 5*, 1067-1072.
458. Keller, D. M., Zeng, X., Wang, Y., Zhang, Q. H., Kapoor, M., Shu, H., Goodman, R., Lozano, G., Zhao, Y., and Lu, H. (2001) A DNA damage-induced p53 serine 392 kinase complex contains CK2, hSpt16, and SSRP1, *Mol. Cell 7*, 283-292.
459. Belotserkovskaya, R., and Reinberg, D. (2004) Facts about FACT and transcript elongation through chromatin, *Curr. Opin. Genet. Dev 14*, 139-146.
460. Blanchoin, L., Boujemaa-Paterski, R., Sykes, C., and Plastino, J. (2014) Actin dynamics, architecture, and mechanics in cell motility, *Physiol Rev 94*, 235-263.
461. Paulin, D., and Li, Z. (2004) Desmin: a major intermediate filament protein essential for the structural integrity and function of muscle, *Exp. Cell Res 301*, 1-7.
462. Dave, J. M., and Bayless, K. J. (2014) Vimentin as an integral regulator of cell adhesion and endothelial sprouting, *Microcirculation 21*, 333-344.
463. Sasiadek, M., Norppa, H., and Sorsa, M. (1991) 1,3-Butadiene and its epoxides induce sister-chromatid exchanges in human lymphocytes in vitro, *Mutat. Res 261*, 117-121.
464. Himmelstein, M. W., Turner, M. J., Asgharian, B., and Bond, J. A. (1996) Metabolism of 1,3-butadiene: inhalation pharmacokinetics and tissue dosimetry of butadiene epoxides in rats and mice, *Toxicology 113*, 306-309.
465. Liu, Y., Fiskum, G., and Schubert, D. (2002) Generation of reactive oxygen species by the mitochondrial electron transport chain, *J. Neurochem 80*, 780-787.
466. Murphy, M. P. (2009) How mitochondria produce reactive oxygen species, *Biochem. J 417*, 1-13.

467. Ide, H., Shoulkamy, M. I., Nakano, T., Miyamoto-Matsubara, M., and Salem, A. M. (2011) Repair and biochemical effects of DNA-protein crosslinks, *Mutat. Res 711*, 113-122.
468. Smith, S. A., Mammen, P. P., Mitchell, J. H., and Garry, M. G. (2003) Role of the exercise pressor reflex in rats with dilated cardiomyopathy, *Circulation 108*, 1126-1132.
469. Louch, W. E., Sheehan, K. A., and Wolska, B. M. (2011) Methods in cardiomyocyte isolation, culture, and gene transfer, *J. Mol. Cell Cardiol 51*, 288-298.
470. Apple, F. S., Murakami, M. M., Christenson, R. H., Campbell, J. L., Miller, C. J., Hock, K. G., and Scott, M. G. (2004) Analytical performance of the i-STAT cardiac troponin I assay, *Clin Chim Acta 345*, 123-127.
471. Apple, F. S., Ler, R., Chung, A. Y., Berger, M. J., and Murakami, M. M. (2006) Point-of-care i-STAT cardiac troponin I for assessment of patients with symptoms suggestive of acute coronary syndrome, *Clin Chem 52*, 322-325.
472. O'Donovan, C., Martin, M. J., Gattiker, A., Gasteiger, E., Bairoch, A., and Apweiler, R. (2002) High-quality protein knowledge resource: SWISS-PROT and TrEMBL, *Brief. Bioinform 3*, 275-284.
473. de Hoon, M. J., Imoto, S., Nolan, J., and Miyano, S. (2004) Open source clustering software, *Bioinformatics 20*, 1453-1454.
474. Zagorski, J., Gellar, M. A., Obraztsova, M., Kline, J. A., and Watts, J. A. (2007) Inhibition of CINC-1 decreases right ventricular damage caused by experimental pulmonary embolism in rats, *J Immunol 179*, 7820-7826.
475. Dries, J. L., Kent, S. D., and Virag, J. A. (2011) Intramyocardial administration of chimeric ephrinA1-Fc promotes tissue salvage following myocardial infarction in mice, *J Physiol 589*, 1725-1740.
476. Daubert, M. A., and Jeremias, A. (2010) The utility of troponin measurement to detect myocardial infarction: review of the current findings, *Vasc. Health Risk Manag 6*, 691-699.
477. Al Saleh, A., Alazzoni, A., Al Shalash, S., Ye, C., Mbuagbaw, L., Thabane, L., and Jolly, S. S. (2014) Performance of the high-sensitivity troponin assay in diagnosing acute myocardial infarction: systematic review and meta-analysis, *CMAJ. Open 2*, E199-E207.
478. Fukuhara, K., Nakashima, T., Abe, M., Masuda, T., Hamada, H., Iwamoto, H., Fujitaka, K., Kohno, N., and Hattori, N. (2017) Suplatast tosilate protects the lung against hyperoxic lung injury by scavenging hydroxyl radicals, *Free Radic. Biol. Med 106*, 1-9.
479. Friedenreich, C. M., Pialoux, V., Wang, Q., Shaw, E., Brenner, D. R., Waltz, X., Conroy, S. M., Johnson, R., Woolcott, C. G., Poulin, M. J., and Courneya, K. S. (2016) Effects of exercise on markers of oxidative stress: an Ancillary analysis of the Alberta Physical Activity and Breast Cancer Prevention Trial, *BMJ Open. Sport Exerc. Med 2*, e000171.
480. Arany, I., Hall, S., Reed, D. K., Reed, C. T., and Dixit, M. (2016) Nicotine Enhances High-Fat Diet-Induced Oxidative Stress in the Kidney, *Nicotine. Tob. Res 18*, 1628-1634.

481. Dizdaroglu, M., and Gajewski, E. (1989) Structure and mechanism of hydroxyl radical-induced formation of a DNA-protein cross-link involving thymine and lysine in nucleohistone, *Cancer Res 49*, 3463-3467.
482. Margolis, S. A., Coxon, B., Gajewski, E., and Dizdaroglu, M. (1988) Structure of a hydroxyl radical induced cross-link of thymine and tyrosine, *Biochemistry 27*, 6353-6359.
483. Charlton, T. S., Ingelse, B. A., Black, D. S., Craig, D. C., Mason, K. E., and Duncan, M. W. (1999) A covalent thymine-tyrosine adduct involved in DNA-protein crosslinks: synthesis, characterization, and quantification, *Free Radic Biol Med 27*, 254-261.
484. Thompson, A., Schafer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A. K., and Hamon, C. (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS, *Anal. Chem 75*, 1895-1904.
485. Lemeer, S., Hahne, H., Pachl, F., and Kuster, B. (2012) Software tools for MS-based quantitative proteomics: a brief overview, *Methods Mol Biol 893*, 489-499.
486. Zweier, J. L., Flaherty, J. T., and Weisfeldt, M. L. (1987) Direct measurement of free radical generation following reperfusion of ischemic myocardium, *Proc. Natl. Acad. Sci. U. S. A 84*, 1404-1407.
487. Xu, Z., Alloush, J., Beck, E., and Weisleder, N. (2014) A murine model of myocardial ischemia-reperfusion injury through ligation of the left anterior descending artery, *J Vis Exp*.
488. Altman, S. A., Zastawny, T. H., Randers-Eichhorn, L., Cacciuttolo, M. A., Akman, S. A., Dizdaroglu, M., and Rao, G. (1995) Formation of DNA-protein cross-links in cultured mammalian cells upon treatment with iron ions, *Free Radic Biol Med 19*, 897-902.
489. Halliwell, B. (2003) Oxidative stress in cell culture: an under-appreciated problem?, *FEBS Lett 540*, 3-6.
490. Kalogeris, T., Bao, Y., and Korthuis, R. J. (2014) Mitochondrial reactive oxygen species: a double edged sword in ischemia/reperfusion vs preconditioning, *Redox. Biol 2*, 702-714.
491. Chen, Y. R., and Zweier, J. L. (2014) Cardiac mitochondria and reactive oxygen species generation, *Circ. Res 114*, 524-537.
492. Lenaz, G. (2012) Mitochondria and reactive oxygen species. Which role in physiology and pathology?, *Adv. Exp. Med. Biol 942*, 93-136.
493. Adam-Vizi, V., and Tretter, L. (2013) The role of mitochondrial dehydrogenases in the generation of oxidative stress, *Neurochem. Int 62*, 757-763.
494. Quinlan, C. L., Goncalves, R. L., Hey-Mogensen, M., Yadava, N., Bunik, V. I., and Brand, M. D. (2014) The 2-oxoacid dehydrogenase complexes in mitochondria can produce superoxide/hydrogen peroxide at much higher rates than complex I, *J. Biol. Chem 289*, 8312-8325.
495. Sutendra, G., Kinnaird, A., Dromparis, P., Paulin, R., Stenson, T. H., Haromy, A., Hashimoto, K., Zhang, N., Flaim, E., and Michelakis, E. D. (2014) A nuclear pyruvate dehydrogenase complex is important for the generation of acetyl-CoA and histone acetylation, *Cell 158*, 84-97.

496. Yim, M. B., Chock, P. B., and Stadtman, E. R. (1990) Copper, zinc superoxide dismutase catalyzes hydroxyl radical production from hydrogen peroxide, *Proc. Natl. Acad. Sci. U. S. A 87*, 5006-5010.
497. Goyal, M. M., and Basak, A. (2012) Hydroxyl radical generation theory: a possible explanation of unexplained actions of mammalian catalase, *Int. J. Biochem. Mol. Biol 3*, 282-289.
498. Cadet, J., Delatour, T., Douki, T., Gasparutto, D., Pouget, J. P., Ravanat, J. L., and Sauvaigo, S. (1999) Hydroxyl radicals and DNA base damage, *Mutat. Res 424*, 9-21.
499. Remppis, A., Ehlermann, P., Giannitsis, E., Greten, T., Most, P., Muller-Bardorff, M., and Katus, H. A. (2000) Cardiac troponin T levels at 96 hours reflect myocardial infarct size: a pathoanatomical study, *Cardiology 93*, 249-253.
500. Metzler, B., Hammerer-Lercher, A., Jehle, J., Dietrich, H., Pachinger, O., Xu, Q., and Mair, J. (2002) Plasma cardiac troponin T closely correlates with infarct size in a mouse model of acute myocardial infarction, *Clin Chim Acta 325*, 87-90.
501. De Celle, T., Vanrobaeys, F., Lijnen, P., Blankesteijn, W. M., Heeneman, S., Van Beeumen, J., Devreese, B., Smits, J. F., and Janssen, B. J. (2005) Alterations in mouse cardiac proteome after in vivo myocardial infarction: permanent ischaemia versus ischaemia-reperfusion, *Exp. Physiol 90*, 593-606.
502. Van Eyk, J. E., Powers, F., Law, W., Larue, C., Hodges, R. S., and Solaro, R. J. (1998) Breakdown and release of myofilament proteins during ischemia and ischemia/reperfusion in rat hearts: identification of degradation products and effects on the pCa-force relation, *Circ Res 82*, 261-271.
503. Liu, T., Chen, L., Kim, E., Tran, D., Phinney, B. S., and Knowlton, A. A. (2014) Mitochondrial proteome remodeling in ischemic heart failure, *Life Sci 101*, 27-36.
504. (2000) Effect of 48-h intravenous trimetazidine on short- and long-term outcomes of patients with acute myocardial infarction, with and without thrombolytic therapy; A double-blind, placebo-controlled, randomized trial. The EMIP-FR Group. European Myocardial Infarction Project--Free Radicals, *Eur. Heart J 21*, 1537-1546.
505. Smith, R. A., and Murphy, M. P. (2011) Mitochondria-targeted antioxidants as therapies, *Discov. Med 11*, 106-114.
506. Starkov, A. A., and Fiskum, G. (2003) Regulation of brain mitochondrial H2O2 production by membrane potential and NAD(P)H redox state, *J. Neurochem 86*, 1101-1107.
507. Logan, A., Cocheme, H. M., Li Pun, P. B., Apostolova, N., Smith, R. A., Larsen, L., Larsen, D. S., James, A. M., Fearnley, I. M., Rogatti, S., Prime, T. A., Finichiu, P. G., Dare, A., Chouchani, E. T., Pell, V. R., Methner, C., Quin, C., McQuaker, S. J., Krieg, T., Hartley, R. C., and Murphy, M. P. (2014) Using exomarkers to assess mitochondrial reactive species in vivo, *Biochim. Biophys. Acta 1840*, 923-930.
508. Kaludercic, N., Deshwal, S., and Di Lisa, F. (2014) Reactive oxygen species and redox compartmentalization, *Front Physiol 5*, 285.
509. Sung, J. S., and Demple, B. (2006) Roles of base excision repair subpathways in correcting oxidized abasic sites in DNA, *FEBS J 273*, 1620-1629.

510. Blaisdell, J. O., and Wallace, S. S. (2007) Rapid determination of the active fraction of DNA repair glycosylases: a novel fluorescence assay for trapped intermediates, *Nucleic Acids Res 35*, 1601-1611.
511. Salem, A. M., Nakano, T., Takuwa, M., Matoba, N., Tsuboi, T., Terato, H., Yamamoto, K., Yamada, M., Nohmi, T., and Ide, H. (2009) Genetic analysis of repair and damage tolerance mechanisms for DNA-protein cross-links in Escherichia coli, *J Bacteriol 191*, 5657-5668.
512. Vaz, B., Popovic, M., and Ramadan, K. (2017) DNA-Protein Crosslink Proteolysis Repair, *Trends Biochem Sci 42*, 483-495.
513. Vaz, B., Popovic, M., Newman, J. A., Fielden, J., Aitkenhead, H., Halder, S., Singh, A. N., Vendrell, I., Fischer, R., Torrecilla, I., Drobnitzky, N., Freire, R., Amor, D. J., Lockhart, P. J., Kessler, B. M., McKenna, G. W., Gileadi, O., and Ramadan, K. (2016) Metalloprotease SPRTN/DVC1 Orchestrates Replication-Coupled DNA-Protein Crosslink Repair, *Mol. Cell 64*, 704-719.
514. Maskey, R. S., Flatten, K. S., Sieben, C. J., Peterson, K. L., Baker, D. J., Nam, H. J., Kim, M. S., Smyrk, T. C., Kojima, Y., Machida, Y., Santiago, A., van Deursen, J. M., Kaufmann, S. H., and Machida, Y. J. (2017) Spartan deficiency causes accumulation of Topoisomerase 1 cleavage complexes and tumorigenesis, *Nucleic Acids Res 45*, 4564-4576.
515. Desai, S. D., Liu, L. F., Vazquez-Abad, D., and D'Arpa, P. (1997) Ubiquitin-dependent destruction of topoisomerase I is stimulated by the antitumor drug camptothecin, *J Biol Chem 272*, 24159-24164.
516. Desai, S. D., Li, T. K., Rodriguez-Bauman, A., Rubin, E. H., and Liu, L. F. (2001) Ubiquitin/26S proteasome-mediated degradation of topoisomerase I as a resistance mechanism to camptothecin in tumor cells, *Cancer Res 61*, 5926-5932.
517. Mao, Y., Desai, S. D., Ting, C. Y., Hwang, J., and Liu, L. F. (2001) 26 S proteasome-mediated degradation of topoisomerase II cleavable complexes, *J Biol Chem 276*, 40652-40658.
518. Zhang, A., Lyu, Y. L., Lin, C. P., Zhou, N., Azarova, A. M., Wood, L. M., and Liu, L. F. (2006) A protease pathway for the repair of topoisomerase II-DNA covalent complexes, *J Biol Chem 281*, 35997-36003.
519. Yang, S. W., Burgin, A. B., Jr., Huizenga, B. N., Robertson, C. A., Yao, K. C., and Nash, H. A. (1996) A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases, *Proc Natl Acad Sci U S A 93*, 11534-11539.
520. Interthal, H., Chen, H. J., and Champoux, J. J. (2005) Human Tdp1 cleaves a broad spectrum of substrates, including phosphoamide linkages, *J Biol Chem 280*, 36518-36528.
521. Interthal, H., and Champoux, J. J. (2011) Effects of DNA and protein size on substrate cleavage by human tyrosyl-DNA phosphodiesterase 1, *Biochem J 436*, 559-566.
522. Zecevic, A., Hagan, E., Reynolds, M., Poage, G., Johnston, T., and Zhitkovich, A. (2010) XPA impacts formation but not proteasome-sensitive repair of DNA-protein cross-links induced by chromate, *Mutagenesis 25*, 381-388.
523. Karu, T. (1999) Primary and secondary mechanisms of action of visible to near-IR radiation on cells, *J Photochem Photobiol B 49*, 1-17.

524. Ward, J. F. (1988) DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability, *Prog Nucleic Acid Res Mol Biol 35*, 95-125.
525. Groehler, A., Degner, A., and Tretyakova, N. Y. (2016) Mass Spectrometry-Based Tools to Characterize DNA-Protein Cross-Linking by Bis-Electrophiles, *Basic Clin. Pharmacol. Toxicol*.
526. Gao, R., Schellenberg, M. J., Huang, S. Y., Abdelmalak, M., Marchand, C., Nitiss, K. C., Nitiss, J. L., Williams, R. S., and Pommier, Y. (2014) Proteolytic degradation of topoisomerase II (Top2) enables the processing of Top2.DNA and Top2.RNA covalent complexes by tyrosyl-DNA-phosphodiesterase 2 (TDP2), *J Biol Chem 289*, 17960-17969.
527. Han, Y. H., Moon, H. J., You, B. R., and Park, W. H. (2009) The effect of MG132, a proteasome inhibitor on HeLa cells in relation to cell growth, reactive oxygen species and GSH, *Oncol Rep 22*, 215-221.
528. Guo, N., and Peng, Z. (2013) MG132, a proteasome inhibitor, induces apoptosis in tumor cells, *Asia Pac J Clin Oncol 9*, 6-11.
529. Fenteany, G., Standaert, R. F., Lane, W. S., Choi, S., Corey, E. J., and Schreiber, S. L. (1995) Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin, *Science 268*, 726-731.
530. Fenteany, G., and Schreiber, S. L. (1998) Lactacystin, proteasome function, and cell fate, *J Biol Chem 273*, 8545-8548.
531. Colvin, M., Brundrett, R. B., Kan, M. N., Jardine, I., and Fenselau, C. (1976) Alkylating properties of phosphoramide mustard, *Cancer Res 36*, 1121-1126.
532. Koivisto, P., Kilpelainen, I., Rasanen, I., Adler, I. D., Pacchierotti, F., and Peltonen, K. (1999) Butadiene diolepoxide- and diepoxybutane-derived DNA adducts at N7-guanine: a high occurrence of diolepoxide-derived adducts in mouse lung after 1,3-butadiene exposure, *Carcinogenesis 20*, 1253-1259.
533. Bouayadi, K., Calsou, P., Pedrini, A. M., and Salles, B. (1992) In vitro evolution of cisplatin/DNA monoadducts into diadducts is dependent upon superhelical density, *Biochem Biophys Res Commun 189*, 111-118.
534. Spanswick, V. J., Lowe, H. L., Newton, C., Bingham, J. P., Bagnobianchi, A., Kiakos, K., Craddock, C., Ledermann, J. A., Hochhauser, D., and Hartley, J. A. (2012) Evidence for different mechanisms of 'unhooking' for melphalan and cisplatin-induced DNA interstrand cross-links in vitro and in clinical acquired resistant tumour samples, *BMC Cancer 12*, 436.
535. Poklar, N., Pilch, D. S., Lippard, S. J., Redding, E. A., Dunham, S. U., and Breslauer, K. J. (1996) Influence of cisplatin intrastrand crosslinking on the conformation, thermal stability, and energetics of a 20-mer DNA duplex, *Proc Natl Acad Sci U S A 93*, 7606-7611.
536. Huang, H., Zhu, L., Reid, B. R., Drobny, G. P., and Hopkins, P. B. (1995) Solution structure of a cisplatin-induced DNA interstrand cross-link, *Science 270*, 1842-1845.
537. Park, S., and Tretyakova, N. (2004) Structural characterization of the major DNA-DNA cross-link of 1,2,3,4-diepoxybutane, *Chem. Res. Toxicol 17*, 129-136.
538. Dong, Q., Barsky, D., Colvin, M. E., Melius, C. F., Ludeman, S. M., Moravek, J. F., Colvin, O. M., Bigner, D. D., Modrich, P., and Friedman, H. S. (1995) A

structural basis for a phosphoramide mustard-induced DNA interstrand cross-link at 5'-d(GAC), *Proc Natl Acad Sci U S A 92*, 12170-12174.
539. Machida, Y., Kim, M. S., and Machida, Y. J. (2012) Spartan/C1orf124 is important to prevent UV-induced mutagenesis, *Cell Cycle 11*, 3395-3402.
540. Kim, M. S., Machida, Y., Vashisht, A. A., Wohlschlegel, J. A., Pang, Y. P., and Machida, Y. J. (2013) Regulation of error-prone translesion synthesis by Spartan/C1orf124, *Nucleic Acids Res 41*, 1661-1668.
541. Leutner, S., Eckert, A., and Muller, W. E. (2001) ROS generation, lipid peroxidation and antioxidant enzyme activities in the aging brain, *J Neural Transm (Vienna) 108*, 955-967.
542. Shah, S. V. (1981) Light emission by isolated rat glomeruli in response to phorbol myristate acetate, *J Lab Clin Med 98*, 46-57.
543. Bromberg, Y., and Pick, E. (1983) Unsaturated fatty acids as second messengers of superoxide generation by macrophages, *Cell Immunol 79*, 240-252.
544. Davies, M., Coles, G. A., and Harber, M. J. (1984) Effect of glomerular basement membrane on the initiation of chemiluminescence and lysosomal enzyme release in human polymorphonuclear leucocytes: an in vitro model of glomerular disease, *Immunology 52*, 151-159.
545. Ferrington, D. A., and Gregerson, D. S. (2012) Immunoproteasomes: structure, function, and antigen presentation, *Prog Mol Biol Transl Sci 109*, 75-112.
546. Neefjes, J., Jongsma, M. L., Paul, P., and Bakke, O. (2011) Towards a systems understanding of MHC class I and MHC class II antigen presentation, *Nat Rev Immunol 11*, 823-836.
547. York, I. A., Goldberg, A. L., Mo, X. Y., and Rock, K. L. (1999) Proteolysis and class I major histocompatibility complex antigen presentation, *Immunol Rev 172*, 49-66.
548. Seifert, U., Bialy, L. P., Ebstein, F., Bech-Otschir, D., Voigt, A., Schroter, F., Prozorovski, T., Lange, N., Steffen, J., Rieger, M., Kuckelkorn, U., Aktas, O., Kloetzel, P. M., and Kruger, E. (2010) Immunoproteasomes preserve protein homeostasis upon interferon-induced oxidative stress, *Cell 142*, 613-624.
549. Hayashi, T., Horiuchi, A., Sano, K., Hiraoka, N., Kanai, Y., Shiozawa, T., Tonegawa, S., and Konishi, I. (2011) Molecular Approach to Uterine Leiomyosarcoma: LMP2-Deficient Mice as an Animal Model of Spontaneous Uterine Leiomyosarcoma, *Sarcoma 2011*, 476498.
550. Ferrington, D. A., Hussong, S. A., Roehrich, H., Kapphahn, R. J., Kavanaugh, S. M., Heuss, N. D., and Gregerson, D. S. (2008) Immunoproteasome responds to injury in the retina and brain, *J Neurochem 106*, 158-169.
551. Hussong, S. A., Kapphahn, R. J., Phillips, S. L., Maldonado, M., and Ferrington, D. A. (2010) Immunoproteasome deficiency alters retinal proteasome's response to stress, *J Neurochem 113*, 1481-1490.
552. Hussong, S. A., Roehrich, H., Kapphahn, R. J., Maldonado, M., Pardue, M. T., and Ferrington, D. A. (2011) A novel role for the immunoproteasome in retinal function, *Invest Ophthalmol Vis Sci 52*, 714-723.
553. Hu, J., Klein, J. D., Du, J., and Wang, X. H. (2008) Cardiac muscle protein catabolism in diabetes mellitus: activation of the ubiquitin-proteasome system by insulin deficiency, *Endocrinology 149*, 5384-5390.

554. Ding, Q., Martin, S., Dimayuga, E., Bruce-Keller, A. J., and Keller, J. N. (2006) LMP2 knock-out mice have reduced proteasome activities and increased levels of oxidatively damaged proteins, *Antioxid Redox Signal 8*, 130-135.
555. Chesner, L. C., C. (2018) A quantitative PCR-based assay reveals that nucleotide excision repair plays a predominant role in the removal of DNA-protein crosslinks from plasmids transfected into mammalian cells, *DNA Repair 62*, 18 - 27.
556. Hecht, S. S. (2003) Tobacco carcinogens, their biomarkers and tobacco-induced cancer, *Nat. Rev. Cancer 3*, 733-744.
557. Gustafson, P., Barregard, L., Strandberg, B., and Sallsten, G. (2007) The impact of domestic wood burning on personal, indoor and outdoor levels of 1,3-butadiene, benzene, formaldehyde and acetaldehyde, *J. Environ. Monit 9*, 23-32.
558. Blair, I. A., Oe, T., Kambouris, S., and Chaudhary, A. K. (2000) 1,3-butadiene: cancer, mutations, and adducts. Part IV: Molecular dosimetry of 1,3-butadiene, *Res. Rep. Health Eff. Inst*, 151-190.
559. Kotapati, S., Wickramaratne, S., Esades, A., Boldry, E. J., Quirk, D. D., Pence, M. G., Guengerich, F. P., and Tretyakova, N. Y. (2015) Polymerase Bypass of N(6)-Deoxyadenosine Adducts Derived from Epoxide Metabolites of 1,3-Butadiene, *Chem. Res. Toxicol 28*, 1496-1507.
560. Carmical, J. R., Zhang, M., Nechev, L., Harris, C. M., Harris, T. M., and Lloyd, R. S. (2000) Mutagenic potential of guanine N(2) adducts of butadiene mono- and diolepoxide volume 13, number 1, january 2000, pp 18-25, *Chem. Res. Toxicol 13*, 430.
561. Steen, A. M., Meyer, K. G., and Recio, L. (1997) Analysis of hprt mutations occurring in human TK6 lymphoblastoid cells following exposure to 1,2,3,4-diepoxybutane, *Mutagenesis 12*, 61-67.
562. Seo, K. Y., Jelinsky, S. A., and Loechler, E. L. (2000) Factors that influence the mutagenic patterns of DNA adducts from chemical carcinogens, *Mutat. Res 463*, 215-246.
563. Tretyakova, N. Y., Chiang, S. Y., Walker, V. E., and Swenberg, J. A. (1998) Quantitative analysis of 1,3-butadiene-induced DNA adducts in vivo and in vitro using liquid chromatography electrospray ionization tandem mass spectrometry, *J. Mass Spectrom 33*, 363-376.
564. Zhao, C., Vodicka, P., RJ, S., and Hemminki, K. (2000) Human DNA adducts of 1,3-butadiene, an important environmental carcinogen, *Carcinogenesis 21*, 107-111.
565. Goggin, M., Swenberg, J. A., Walker, V. E., and Tretyakova, N. (2009) Molecular dosimetry of 1,2,3,4-diepoxybutane-induced DNA-DNA cross-links in B6C3F1 mice and F344 rats exposed to 1,3-butadiene by inhalation, *Cancer Res 69*, 2479-2486.
566. Goggin, M., Anderson, C., Park, S., Swenberg, J., Walker, V., and Tretyakova, N. (2008) Quantitative high-performance liquid chromatography-electrospray ionization-tandem mass spectrometry analysis of the adenine-guanine cross-links of 1,2,3,4-diepoxybutane in tissues of butadiene-exposed B6C3F1 mice, *Chem. Res. Toxicol 21*, 1163-1170.

567. Seneviratne, U., Antsypovich, S., Goggin, M., Dorr, D. Q., Guza, R., Moser, A., Thompson, C., York, D. M., and Tretyakova, N. (2010) Exocyclic deoxyadenosine adducts of 1,2,3,4-diepoxybutane: synthesis, structural elucidation, and mechanistic studies, *Chem. Res. Toxicol 23*, 118-133.
568. Watson, J. D., and Crick, F. H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid, *Nature 171*, 737-738.
569. Christov, P. P., Banerjee, S., Stone, M. P., and Rizzo, C. J. (2010) Selective Incision of the alpha-N-Methyl-Formamidopyrimidine Anomer by Escherichia coli Endonuclease IV, *J Nucleic Acids 2010*.
570. Hu, J., de Souza-Pinto, N. C., Haraguchi, K., Hogue, B. A., Jaruga, P., Greenberg, M. M., Dizdaroglu, M., and Bohr, V. A. (2005) Repair of formamidopyrimidines in DNA involves different glycosylases: role of the OGG1, NTH1, and NEIL1 enzymes, *J Biol Chem 280*, 40544-40551.
571. Dherin, C., Radicella, J. P., Dizdaroglu, M., and Boiteux, S. (1999) Excision of oxidatively damaged DNA bases by the human alpha-hOgg1 protein and the polymorphic alpha-hOgg1(Ser326Cys) protein which is frequently found in human populations, *Nucleic Acids Res 27*, 4001-4007.
572. Chan, M. K., Ocampo-Hafalla, M. T., Vartanian, V., Jaruga, P., Kirkali, G., Koenig, K. L., Brown, S., Lloyd, R. S., Dizdaroglu, M., and Teebor, G. W. (2009) Targeted deletion of the genes encoding NTH1 and NEIL1 DNA N-glycosylases reveals the existence of novel carcinogenic oxidative damage to DNA, *DNA Repair (Amst) 8*, 786-794.
573. Vartanian, V., Minko, I. G., Chawanthayatham, S., Egner, P. A., Lin, Y. C., Earley, L. F., Makar, R., Eng, J. R., Camp, M. T., Li, L., Stone, M. P., Lasarev, M. R., Groopman, J. D., Croy, R. G., Essigmann, J. M., McCullough, A. K., and Lloyd, R. S. (2017) NEIL1 protects against aflatoxin-induced hepatocellular carcinoma in mice, *Proc Natl Acad Sci U S A 114*, 4207-4212.
574. Klich, M. A. (2007) Aspergillus flavus: the major producer of aflatoxin, *Mol Plant Pathol 8*, 713-722.
575. Schindler, A. F., Palmer, J. G., and Eisenberg, W. V. (1967) Aflatoxin Production by Aspergillus flavus as Related to Various Temperatures, *Appl Microbiol 15*, 1006-1009.
576. Li, F. Q., Yoshizawa, T., Kawamura, O., Luo, X. Y., and Li, Y. W. (2001) Aflatoxins and fumonisins in corn from the high-incidence area for human hepatocellular carcinoma in Guangxi, China, *J Agric Food Chem 49*, 4122-4126.
577. Ding, X., Wu, L., Li, P., Zhang, Z., Zhou, H., Bai, Y., Chen, X., and Jiang, J. (2015) Risk Assessment on Dietary Exposure to Aflatoxin B(1) in Post-Harvest Peanuts in the Yangtze River Ecological Region, *Toxins (Basel) 7*, 4157-4174.
578. Shimada, T., and Guengerich, F. P. (1989) Evidence for cytochrome P-450NF, the nifedipine oxidase, being the principal enzyme involved in the bioactivation of aflatoxins in human liver, *Proc. Natl. Acad. Sci. U. S. A 86*, 462-465.
579. Hertzog, P. J., Smith, J. R., and Garner, R. C. (1982) Characterisation of the imidazole ring-opened forms of trans-8,9-dihydro-8,9-dihydro-8-(7-guanyl)9-hydroxy aflatoxin B1, *Carcinogenesis 3*, 723-725.
580. Gruppi, F., Hejazi, L., Christov, P. P., Krishnamachari, S., Turesky, R. J., and Rizzo, C. J. (2015) Characterization of nitrogen mustard formamidopyrimidine adduct

formation of bis(2-chloroethyl)ethylamine with calf thymus DNA and a human mammary cancer cell line, *Chem Res Toxicol 28*, 1850-1860.
581. Vartanian, V., Lowell, B., Minko, I. G., Wood, T. G., Ceci, J. D., George, S., Ballinger, S. W., Corless, C. L., McCullough, A. K., and Lloyd, R. S. (2006) The metabolic syndrome resulting from a knockout of the NEIL1 DNA glycosylase, *Proc Natl Acad Sci U S A 103*, 1864-1869.
582. CAVALIERI, L. F., BENDICH, A., and . (1948) Ultraviolet absorption spectra of purines, pyrimidines and triazolopyrimidines, *J. Am. Chem. Soc 70*, 3875-3880.
583. Dizdaroglu, M., Kirkali, G., and Jaruga, P. (2008) Formamidopyrimidines in DNA: mechanisms of formation, repair, and biological effects, *Free Radic Biol Med 45*, 1610-1621.
584. Rusyn, I., Asakura, S., Li, Y., Kosyk, O., Koc, H., Nakamura, J., Upton, P. B., and Swenberg, J. A. (2005) Effects of ethylene oxide and ethylene inhalation on DNA adducts, apurinic/apyrimidinic sites and expression of base excision DNA repair genes in rat brain, spleen, and liver, *DNA Repair (Amst) 4*, 1099-1110.
585. Rios-Blanco, M. N., Faller, T. H., Nakamura, J., Kessler, W., Kreuzer, P. E., Ranasinghe, A., Filser, J. G., and Swenberg, J. A. (2000) Quantitation of DNA and hemoglobin adducts and apurinic/apyrimidinic sites in tissues of F344 rats exposed to propylene oxide by inhalation, *Carcinogenesis 21*, 2011-2018.
586. O'Brien, P. J., and Ellenberger, T. (2004) Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase, *J Biol Chem 279*, 9750-9757.
587. Berger, M., and Cadet, J. (1985) Isolation and Characterization of the Radiation-Induced Degradation Products of 2'-Deoxyguanosine in Oxygen-Free Aqueous-Solutions, *Zeitschrift Fur Naturforschung Section B-a Journal of Chemical Sciences 40*, 1519-1531.
588. Boiteux, S., and Laval, J. (1983) Imidazole open ring 7-methylguanine: an inhibitor of DNA synthesis, *Biochem Biophys Res Commun 110*, 552-558.
589. Ji, S., Shao, H., Han, Q., Seiler, C. L., and Tretyakova, N. Y. (2017) Reversible DNA-Protein Cross-Linking at Epigenetic DNA Marks, *Angew Chem Int Ed Engl 56*, 14130-14134.
590. Li, D., Wu, J., Bai, Y., Zhao, X., and Liu, L. (2014) Isolation and culture of adult mouse cardiomyocytes for cell signaling and in vitro cardiac hypertrophy, *J. Vis. Exp*.
591. Chen, M., Yang, Z. W., Zhu, J. T., Xiao, Z. Y., and Xiao, R. (1990) [Anti-arrhythmic effects and electrophysiological properties of Ophiopogon total saponins], *Zhongguo Yao Li Xue Bao 11*, 161-165.
592. Chen, X., Tang, J., Xie, W., Wang, J., Jin, J., Ren, J., Jin, L., and Lu, J. (2013) Protective effect of the polysaccharide from Ophiopogon japonicus on streptozotocin-induced diabetic rats, *Carbohydr Polym 94*, 378-385.
593. Ichikawa, H., Wang, X., and Konishi, T. (2003) Role of component herbs in antioxidant activity of shengmai san--a traditional Chinese medicine formula preventing cerebral oxidative damage in rat, *Am J Chin Med 31*, 509-521.
594. Kou, J., Tian, Y., Tang, Y., Yan, J., and Yu, B. (2006) Antithrombotic activities of aqueous extract from Radix Ophiopogon japonicus and its two constituents, *Biol Pharm Bull 29*, 1267-1270.

595. Qian, J., Jiang, F., Wang, B., Yu, Y., Zhang, X., Yin, Z., and Liu, C. (2010) Ophiopogonin D prevents H2O2-induced injury in primary human umbilical vein endothelial cells, *J Ethnopharmacol 128*, 438-445.
596. Zhang, Y. Y., Meng, C., Zhang, X. M., Yuan, C. H., Wen, M. D., Chen, Z., Dong, D. C., Gao, Y. H., Liu, C., and Zhang, Z. (2015) Ophiopogonin D attenuates doxorubicin-induced autophagic cell death by relieving mitochondrial damage in vitro and in vivo, *J Pharmacol Exp Ther 352*, 166-174.
597. Huang, Y. L., Kou, J. P., Ma, L., Song, J. X., and Yu, B. Y. (2008) Possible mechanism of the anti-inflammatory activity of ruscogenin: role of intercellular adhesion molecule-1 and nuclear factor-kappaB, *J Pharmacol Sci 108*, 198-205.
598. Xie, T., Liang, Y., Hao, H., A, J., Xie, L., Gong, P., Dai, C., Liu, L., Kang, A., Zheng, X., and Wang, G. (2012) Rapid identification of ophiopogonins and ophiopogonones in Ophiopogon japonicus extract with a practical technique of mass defect filtering based on high resolution mass spectrometry, *J Chromatogr A 1227*, 234-244.
599. Klier, L. Z., D.; Rahimoff,R.; Mosrin,M.; Knochel,P. (2017) Practical Large-Scale Regioselective Zincation of Chromone Using TMPZnCl·LiCl Triggered by the Presence or Absence of MgCl2, *Org. Process. Res. Dev. 21*, 660 - 663.
600. Jeganmohan, M., and Knochel, P. (2010) tmp(4)Zr: an atom-economical base for the metalation of functionalized arenes and heteroarenes, *Angew Chem Int Ed Engl 49*, 8520-8524.
601. Hamid, A. A., Hasanain, M., Singh, A., Bhukya, B., Omprakash, Vasudev, P. G., Sarkar, J., Chanda, D., Khan, F., Aiyelaagbe, O. O., and Negi, A. S. (2014) Synthesis of novel anticancer agents through opening of spiroacetal ring of diosgenin, *Steroids 87*, 108-118.
602. Gray, G. M., and Yardley, H. J. (1975) Mitochondria and nuclei of pig and human epidermis: isolation and lipid composition, *J Invest Dermatol 64*, 423-430.
603. Dimauro, I., Pearson, T., Caporossi, D., and Jackson, M. J. (2012) A simple protocol for the subcellular fractionation of skeletal muscle cells and tissue, *BMC Res Notes 5*, 513.
604. Bennett, B. T., Bewersdorf, J., and Knight, K. L. (2009) Immunofluorescence imaging of DNA damage response proteins: optimizing protocols for super-resolution microscopy, *Methods 48*, 63-71.
605. Nolan, T., Hands, R. E., and Bustin, S. A. (2006) Quantification of mRNA using real-time RT-PCR, *Nat Protoc 1*, 1559-1582.
606. Chen, H. J., Hsieh, C. J., Shen, L. C., and Chang, C. M. (2007) Characterization of DNA--protein cross-links induced by oxanine: cellular damage derived from nitric oxide and nitrous acid, *Biochemistry 46*, 3952-3965.
607. Marcos, J., and Pozo, O. J. (2015) Derivatization of steroids in biological samples for GC-MS and LC-MS analyses, *Bioanalysis 7*, 2515-2536.
608. Toribio-Delgado, A. F., Maynar-Marino, M., Caballero-Loscos, M. J., Robles-Gil, M. C., Olcina-Camacho, G. J., and Maynar-Marino, J. I. (2012) Qualification and quantification of seventeen natural steroids in plasma by GC-Q-MS and GC-IT-MS/MS, *J Chromatogr Sci 50*, 349-357.
609. Nordstrom, A., Tarkowski, P., Tarkowska, D., Dolezal, K., Astot, C., Sandberg, G., and Moritz, T. (2004) Derivatization for LC-electrospray ionization-MS: a tool

for improving reversed-phase separation and ESI responses of bases, ribosides, and intact nucleotides, *Anal Chem 76*, 2869-2877.
610. Licea-Perez, H. J., V.; Zohrabian,S.; Karlinsey,M. (2016) Development of a multi-sugar LC-MS/MS assay using simple chemical derivatization with acetic anhydride, *Anal Methods 8*, 3023 - 3033.
611. Zhang, W. M., Liao, S. J., Xu, Y., and Zhang, Y. P. (1997) Application of alkali metal hydrides of nanometric size in reduction, cyclotrimerization and metalation, *Synthetic Commun 27*, 3977-3983.
612. Kalvin, D. M., and Woodard, R. W. (1984) Preparation of 1-Deuterioaldehydes Via the Use of Diisobutylaluminum Deuteride (Dibal-D), *Tetrahedron 40*, 3387-3392.
613. Kim, S., and Ahn, K. H. (1984) Ate Complex from Diisobutylaluminum Hydride and Normal-Butyllithium as a Powerful and Selective Reducing Agent for the Reduction of Selected Organic-Compounds Containing Various Functional-Groups, *J Org Chem 49*, 1717-1724.
614. Li, L., Brown, K. L., Ma, R., and Stone, M. P. (2015) DNA Sequence Modulates Geometrical Isomerism of the trans-8,9- Dihydro-8-(2,6-diamino-4-oxo-3,4-dihydropyrimid-5-yl-formamido)- 9-hydroxy Aflatoxin B1 Adduct, *Chem Res Toxicol 28*, 225-237.
615. Caruthers, M. H., Barone, A. D., Beaucage, S. L., Dodds, D. R., Fisher, E. F., McBride, L. J., Matteucci, M., Stabinsky, Z., and Tang, J. Y. (1987) Chemical synthesis of deoxyoligonucleotides by the phosphoramidite method, *Methods Enzymol 154*, 287-313.
616. Kowal, E. A., Wickramaratne, S., Kotapati, S., Turo, M., Tretyakova, N., and Stone, M. P. (2014) Major groove orientation of the (2S)-N(6)-(2-hydroxy-3-buten-1-yl)-2'-deoxyadenosine DNA adduct induced by 1,2-epoxy-3-butene, *Chem Res Toxicol 27*, 1675-1686.
617. Maddukuri, L., Eoff, R. L., Choi, J. Y., Rizzo, C. J., Guengerich, F. P., and Marnett, L. J. (2010) In vitro bypass of the major malondialdehyde- and base propenal-derived DNA adduct by human Y-family DNA polymerases kappa, iota, and Rev1, *Biochemistry 49*, 8415-8424.
618. Hess, M. T., Gunz, D., Luneva, N., Geacintov, N. E., and Naegeli, H. (1997) Base pair conformation-dependent excision of benzo[a]pyrene diol epoxide-guanine adducts by human nucleotide excision repair enzymes, *Mol Cell Biol 17*, 7069-7076.